SEQUENCE LISTING

<110> WAKITA, Takaji
      KATO, Takanobu
      DATE, Tomoko

<120> A NUCLEIC ACID CONSTRUCT CONTAINING A NUCLEIC ACID DERIVED
      FROM THE GENOME OF HEPATITIS C VIRUS (HCV) OF GENOTYPE 2a,
      AND A CELL HAVING SUCH NUCLEIC ACID CONSTRUCT INTRODUCED THEREIN

<130> 1254-0299PUS1

<140> US 10/558,155
<141> 2005-11-23

<150> PCT/JP2003/015038
<151> 2003-11-25

<150> JP 2003-148242
<151> 2003-05-26

<150> JP 2003-329115

<151> 2003-09-19

<160> 41

<170> PatentIn Ver. 2.1

<210> 1
<211> 8024
<212> RNA
<213> Artificial Sequence

<220>
<223> HCV-derived synthetically created RNA replicon
      rSGREP-JFH1 sequence

<400> 1
accugccccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360
ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420
cgcagguucu ccggccgcuu ggguggagag cuauucggc uaugacuggg cacaacagac 480
aaucggcugc ucugaugccg ccguguuccg cugucagcg caggggcgcc cgguucuuuu 540
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccugccgag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780
ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900
cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960
uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020
cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080

1

```
ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140
ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200
ccucuccuc cccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260
cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320
aaccuggccc ugucuucuug acgagcauuc cuaggggucu uuccccucuc gccaaaggaa 1380
ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440
caacgucugu agcgacccuu ugcaggcagc ggaacccccc accggcgac aggugccucu 1500
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaaccc cagugccacg 1560
uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620
ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680
caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggcccccg aaccacgggg 1740
acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800
caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860
caggccgggg aaguccaaau ccuguccaca gucucucagu ccuuccucgg aacaaccauc 1920
ucgggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980
gguccgguca cgcagaugua cucgagugcu gaggggggacu ugguaggcug gcccagcccc 2040
ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100
cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160
ccgagcccca uuucgaccuu gaagggguccc ucgggggggc cggugcucug cccuagggge 2220
cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280
uucaucccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340
acgccaccgg cuguugcccca gaccuaucag gucgguacu ugcaugcucc aacuggcagu 2400
ggaaagagca ccaaggucuc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460
aaccccucgg uagcugccac ccuggggguu ggggcguacc uauccaaggc acauggcauc 2520
aauccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580
acauauggca aauuucucgc cgauggggc ugcgcuagcg gcgccuauga caucaucaua 2640
ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700
caagcagaga cagccggggu cagacuaacu gugcuggcua cggccacacc ccccgggguca 2760
gugacaaccc cccaucccga uauagaagag guaggccucg ggcgggaggg ugagaucccc 2820
uucuauggga gggcgauucc ccuauccugc aucaagggag ggagacaccu gauuuucugc 2880
cacucaaaga aaaagugu ga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940
guggcauacu auagaggguu ggacgucucc auaauaccag cucaggagaa uggguguguc 3000
gucgccaccg acgcccucau gacgggguac acuggagacu uugacuccgu gaucgacugc 3060
aauguagcgg ucacccaagc ugucgacuuc agccuggacc ccaccuucac uauaaccaca 3120
cagacugucc cacaagacgc ugucucacgc agucagcgcc gcgggcgcac agguagagga 3180
agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240
guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300
accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360
cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucucc 3420
caaacaaagc aagcggggga gaacuucgcg uaccaguag ccuaccaagc uacggugugc 3480
gccagagcca aggcccccucc cccguccugg gacgccaugu ggaagugccu ggcccgacuc 3540
aagccuacgc uugcgggccc cacaccucuc cuguaccguu ugggcccuau uaccaaugag 3600
gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660
gucaugacca gcacguggu cuuagcugga ggaguccugg cagccgucgc cgcauauugc 3720
cuggcgacug gaugcguuuc caucaucggc cgcuugcacg ucaaccagcg agucgucguu 3780
gcgccggaua aggagguccu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840
gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900
uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960
cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020
cucgcaggau ugucaacacu gccaggaaac cccgcggugg cuuccaugau ggcauucagu 4080
gccgcccuca ccaguccguu gucgaccagu accaccaucc uucucaacau cauggggaggc 4140
ugguuagcgu cccagaucgc accaccgcg ggggccaccg gcuuugucgu caguggccug 4200
gugggggcug ccguggggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260
uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320
ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccuggugggug 4380
ggggucaucu gcgcggccau ucugcgccgc cacguggggac cgggggaggg cgcgguccaa 4440
uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500
```

2

```
gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560
agccuacuca gaagacucca caauuggaua acugaggacu gccccauccc augcuccgga 4620
uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680
cugaccucua aauuguuccc caagcugccc ggccuccccu ucaucucuug ucaaaagggg 4740
uacaaggguug uguggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800
aucucuggca auguccgccu gggcucuaug aggaucacag ggccaaaac cugcaugaac 4860
accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcgc gccgaaaccc 4920
cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980
cagcauggu cguacuccua uguaacagga cugaccacug acaaucugaa aauuccuugc 5040
caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100
cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160
gcugucgggu cccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220
cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280
ccuccaucug aggcgagcuc cucagugagc cagcuaucag caccgucgcu gcgggccacc 5340
ugcaccaccc acagcaacac cuaugacgug gacaauggucg augccaaccu gcucauggag 5400
ggcggugugg cucagacaga gccugaguec aggguugcccg uucuggacuu ucucgagcca 5460
auggccgagg aagagagcga ccuugagccc ucaauaccau cggagugcau gcuccccagg 5520
agcgggguuc cacgggccuu accggcuugg gcacggccug acuacaaccc gccgcucgug 5580
gaaucguggag ggaggccaga uuaccaaccg cccaccguug cugguugugc ucuccccccc 5640
cccaagaagg ccccgacgcc ucccccaagg agacgccgga caguggggucu gagcgagagc 5700
accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760
ggugaugcag gcucguccac ggggggcgggc gccgcgaau ccggcggucc gacgucecccu 5820
ggugagccgg cccccucaga gacagguucc gccuccucua ugccccccu cgaggggggag 5880
ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucccec ccaggggggg 5940
gggguagcuc ccgguucggg cucgggguucu uggucuacuu gcuccgagga ggacgauacc 6000
accgugugcu gcuccaugucc auacuccugg accgggggcuc uaauaaccc cuguagcccc 6060
gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180
acgcaagugc ucgacgccca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240
aaggucagcg caaggccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300
gcaagaucca aguauggauu cgggggccaag gagguccgca gcuuguccgg gagggccguu 6360
aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420
accaucaugg ccaaaaauga ggguuucugc guggacccccg ccaagggggg uaagaaacca 6480
gcucgccuca ucguuuacc ugaccucggc guccggggcu gcgagaaaau ggcccucuau 6540
gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600
ccugcccaac gggguggagua ucucuugaaa gcauggggcgg aaaagaagga ccccauggguu 6660
uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720
gaguccauau accaggccug cucccugccc gaggaggccc gcacugccau acacucgcug 6780
acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840
agacguugcc gcgccagcgg ggugcuaacc acuagcaugg uaacaccau cacaugcuau 6900
gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau gcugguaugc 6960
ggcgaugacc uaguagucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020
agagccuuca cggaggccau gaccagguac ucugccccuc cuggugaucc ccccagaccg 7080
gaauaugacc uggagcuaau aacauccugu uccucaaaug ugucuguggc guugggcccg 7140
cggggccgcc gcagauacua ccugaccaga gacccaacca uccacucgc ccgggcugcc 7200
ugggaaacag uuagacacuc cccuaucaau ucauggcugg gaaacaucau ccaguaugcu 7260
ccaaccauau ggguucgcau gguccuaaug acacacuucu ucuccauucu caugguccaa 7320
gacacccugg accagaaccu caacuuugag auguauggau caguauacuc cgugaauccu 7380
uuggaccuuc cagccauaau ugagagguua cacgggcuug acgccuuuuc uaugcacaca 7440
uacucucacc acgaacugac gcggguggcu ucagcccuca gaaaacuugg ggcgccacccc 7500
cucaggguggu ggaagagucg ggcucgcgca gucaggggcgu cccucaucuc ccguggaggg 7560
aaagcggccg uuugcggccg auaucucuuc aauugggcgg ugaagaccaa gcucaaacuc 7620
acuccauugc cggaggcgcg ccuacuggac uuauccaguu gguucaccgu cggcgccggc 7680
ggggggcgaca uuuuucacag cgugucgcgc gcccgacccc gcucauuacu cuucggccua 7740
cuccuacuuu ucguaggggu aggccucuuc cuacucccccg cucgguagag cggcacacac 7800
uagguacacu ccauagcuaa cuguuccuuu uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu 7860
uuuuuuuuuu cuuuuuuuuu uuuuucccuc uuucuucccu ucucaucuua uucuacuuuc 7920
```

3

```
uuucuuggug gcuccaucuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980
augacugcag agagugccgu aacuggucuc ucugcagauc augu              8024


<210> 2
<211> 8024
<212> RNA
<213> Artificial Sequence

<220>
<223> HCV-derived synthetically created RNA Replicon rSGREP-JCH1

<400> 2
acccgccccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60
cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccucccg ggagagccau aguggucugc ggaaccggu6 aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu uggguacug ccugauaggg 300
ugcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa aucccaaacc 360
ucaaagaaaa accaaaagaa acacuaaccg ucgcccaaug auugaacaag auggauugca 420
cgcagguucu ccggccgcuu gguggagag gcuauucggc uaugacuggg cacaacagac 480
aaucggcugc ucugaugccg ccguguuccg cugucagcg caggggcgcc cgguucuuuu 540
ugucaagacc gaccuguccg gugcccugaa ugaacugcag acgaggcag cgcggcuauc 600
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccugccgag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780
ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac uacucggau 840
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucagggc ucgcgccagc 900
cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960
uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020
cguggccgg cuggguugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080
ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140
ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200
ccucucccuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260
cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320
aaccuggccc ugucuucuug acgagcauuc cuaggggucu uuccccucuc gccaaaggaa 1380
ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440
caacgucugu agcgacccuu gcaggcagc ggaaccccc accggcgac aggugccucu 1500
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaaccc cagugccacg 1560
uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620
ggcugaagga ugcccagaag uaccccauu guaugggauc ugaucugggg ccucggugca 1680
caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggcccccccg aaccacgggg 1740
acgugguuuu ccuuugaaaa acacgauaau accauggccc ccaucaccgc uuacgcccag 1800
cagacacgag gucucuuggg cucuauagug gugagcauga cggggcguga caagacagaa 1860
caggccgggg agguccaagu ccuguccaca gucacucagu ccuuccucgg aacauccauu 1920
ucgggggucu uauggacugu uuaccacgga gcuggcaaca agacacuagc cggcucgcgg 1980
ggcccgguca cgcagaugua cucgagcgcc gaggggggacu uggucggguug gcccagcccu 2040
ccugggacca aaacuuugga gccguuacg uguggagcgg ucgaccugua uuuggucacg 2100
cggaacgcug augucauccc ggcucgaaga cgcggggaca agcgggggagc gcugcucucc 2160
ccgagacccc uuucgaccuu gaaggggucc ucggggggac cugugcuuug cccuagggc 2220
cacgcugucg gaaucuuccg ggcagcugug ugcucucggg guguggcuaa guccauagau 2280
uucaucccc uugagacgcu cgacaucguc acgcggcucu ccaccuuuag ugacaacagc 2340
acaccaccag cuguugcccca gaccuaucag gugggguacu ugcacgcccc cacuggcagu 2400
ggaaaaagca ccaagguccc cgucgcguac gccgcccagg gguauaaagu gcggggugcuc 2460
aaucccucgg uggcugccac ccugggauuu ggggcguacu uguccaaggc acauggcauc 2520
aaccccaaca uuaggacugg agucagaacu gugacgaccg gggagcccau uacauacucc 2580
acguaggua aauuccucgc cgauggggg ugcgcaggcg gcgccuauga caucaucaua 2640
```

4

```
ugcgaugaau gccacucugu ggaugcuacc acuauucucg gcaucgggac aguccuugac 2700
caagcagaga cagccggggu caggcuaacu guacuggcca cggccacgcc ccccgggucg 2760
gugacaaccc cccaucccaa uauagaggag guagcccucg dacaggaggg ugagaucccc 2820
uucuauggga gggcguuucc ccugucuuac aucaagggag ggaggcacuu gauuuucugc 2880
cacucaaaga aaaagugaga cgagcucgca acggcccuuc ggggcauggg cuugaacgcu 2940
guggcauauu acagaggguu ggacgucucc auaauaccaa cucaaggaga uguggugguc 3000
guugccaccg acgcccucau gacgggguau acuggagacu uugacuccgu gaucgacugc 3060
aacguagcgg ucacccaggc cguagacuuc agccuggacc ccaccuucac uauaaccaca 3120
cagacugucc cgcaagacgc ugucucacgu agucagcgcc gagggcgcac ggguagagga 3180
agacugggca uuuauaggua uguuuccacu ggugagcgag ccucaggaau guuugacagu 3240
guaguacucu gugagugcua cgacgcagga gcugcuuggu augagcucuc accaguggag 3300
acgaccguca ggcucagggc guauuucaac acgccuggcu ugccugugug ccaggaccac 3360
cuugaguuuu gggaggcagu uuucaccggc cucacacaca uagacgcuca uuuccuuucc 3420
cagacaaagc agucggggga aaauuucgca uacuuaguag ccuaucaggc cacagugugc 3480
gccagggcca aagcgccccc cccguccugg dacgucaugu ggaagugcuu gacucgacuc 3540
aagcccacgc uuguggggccc uacaccucuc cuguaccguu ugggcucugu uaccaacgag 3600
gucacccuua cacacccccgu gacaaaauac aucgccacau gcaugcaagc ugaccucgag 3660
gucaugacca gcacgu9ggu ccuggcuggg ggagucuuag cagccgucgc cgcguauugc 3720
uuagcgaccg ggugugu uuc caucauuggc cguuuacaca ucaaccagcg agcugucguc 3780
gcuccggaca aggagguccu cuaugaggcu uuugaugaga uggaggaaug ugccuccaga 3840
gcggcucucc uugaagaggg gcagcggaua gccgagaugc ugaaguccaa gauccaaggc 3900
uuauugcagc aagccucuaa acaggcccag dacauacaac ccgcugugca agcuucgugg 3960
cccaagaugg agcaauucug ggccaaacau auguggaacu ucauaagcgg cauucaguac 4020
cucgcaggac ugucaacacu gccagggaac ccugcugugg cuuccaugau ggcauucagc 4080
gccgcccuca ccaguccguu gucaacuagc accaccaucc uucuuaacau ucuggggggc 4140
uggcuggcgu cccaaauugc gccacccgcg ggggccacug gcuuuguugu caguggccug 4200
gugggagcug cuguuggcag cauaggcuug gguaaagugc ugguggacau ccuggcaggg 4260
uauggugcgg gcauuucggg ggcccucguc gcguuaaga ucaugucugg cgagaagccc 4320
uccauggagg augucaucaa cuugcugccu gggauucugu uccaggugc ucugguggug 4380
ggagucaucu gcgcggccau ucugcgccgc caugugggac cgggggaagg cgcgguccaa 4440
uggaugaaca ggcuuaucgc cuucgcuucc agaggaaacc acgucgcccc uacucacuac 4500
gugacggagu cggaugcguc gcagcguguc acccaacugc uuggcucucu cacuauaaacu 4560
agucuacuca ggagacuuca caacuggauc acugaggauu gccccauccc augcgccggc 4620
ucguggcucc gcgaugugug ggacugggu uguaccaucc uaacagacuu uaagaacugg 4680
cugaccucca agcuguuccc aaagaugccu ggccuc.cccu uuaucucuug ccaaaagggg 4740
uacaagggcg ugugggccgg cacuggcauc augaccacac gaugccccug cggcgccaac 4800
aucucuggca acguccgcuu gggcucuaug agaaucacag acccaaaac cugcaugaac 4860
accuggcagg ggaccuuucc uaucaauugu uauacagaag gccagugcuu gccgaaaccc 4920
gcguuaaacu ucaagaccgc caucuggaga guggcggccu cagaguacgc ggaagugacg 4980
cagcacggau cauaugccua uauaacaggg cugaccacug acaacuuaaa aguccuugc 5040
caacuccccu cuccagaguu uuucucuugg guggacggag uacaaaucca uaggccgcc 5100
cccacaccaa agccguuuuu ccgggaugag gucucguuca gcguugggcu caauucauuu 5160
gucgucgggu cucagcuucc cugugaccuu gagcccgaca cugagguagu gauguccaug 5220
cuaacagacc cauccccauau cacggcggag gcugcagcgc ggcguuagc gcggggguca 5280
cccccaucug aggcaagcuc cucagcgagc cagcugucgg cgccaucgcu gcgagccacc 5340
ugcaccaccc acgguaggac cuaugaugug gacauggugg augccaaccu guucaugggg 5400
ggcggcguga uucggauaga gucugaagucc aaaguggucg uucuggacuc ccucgacuca 5460
augaccgagg aagagggcga ccuugagccu ucaguaccau cggaguauau gcuccccagg 5520
aagagguucc caccggccuu accggcuugg gcgcggccug auuacaaccc accgcuugug 5580
gaaucgugga agaggccaga uuaccaacca cccacuguug cgggcugugc ucuccccccc 5640
cccaaaaaga ccccgacgcc uccuccaagg agacgccgga caguggguu gagcgagagc 5700
accauaggag augcccucca acagcuggcc aucaagccu uggccagcc ccccccaagc 5760
ggcgauucag gccuuuccac ggggcggac gccgccgacu ccggcgaucg gacacccccu 5820
gacgaguugg cucuuucgga gacagguucu accuccucca ugcccccccu cgaggggag 5880
ccuggggacc cagaccugga gccugagcag guagagcuuc aaccuccucc cagggggggg 5940
gaggcagcuc ccggcucgga cucgggguccc ugguccuuacuu gcuccgagga ggaugacccc 6000
gucgugugcu gcuccaaguc auauuccugg accggggcuc uaauaacucc uuguagcccc 6060
```

```
gaagaggaaa aguugccaau uaacuccuug agcaacucgc uguugcgaua ccauaacaag 6120
guauacugua cuacaucaaa gagugccuca cuaaggggcua aaaagguaac uuuugauagg 6180
augcaagugc ucgacgccua uuaugauuca gucuuaaagg acaucaagcu agcggccucc 6240
aaggucagcg caaggcuccu caccuuagag gaggcgugcc aauugacccc accccacucu 6300
gcaagaucca aguauggguu uggggcuaag gagguccgca gcuuguccgg gagggccguc 6360
aaccacauca aguccgugug gaaggaccuc uuggaagacu cacaaacacc aauuccuaca 6420
accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaaaaaacca 6480
gcucgccuua ucguuuaccc ugaccucggc gucaggggucu gcgagaagau ggcccuuuau 6540
gaugucacac aaaagcuucc ucaggcggug augggggcuu cuuauggcuu ccaguacucc 6600
cccgcucagc ggguggaguu ucucuugaag gcaugggcgg aaaagagaga cccuauggguu 6660
uuuucguaug auacccgaug cuuugacuca accgucacug agagagacau caggacugag 6720
gaguccauau accaggccug cuccuuaccc gaggaggccc gaacugccau acacucgcug 6780
acugagagac ucuauguggg aggggcccaug uucaacagca agggccaguc cugcggguac 6840
aggcguugcc gcgccagcgg ggugcuuacc acuaguaugg ggaacaccau cacaugcuau 6900
guaaaagccc uagcggcuug caaggcugcg gggauaauug cgcccacgau gcugguaugc 6960
ggcgacgacu ggucgucauc ucagaaaagc caggggacug aggaggacga gcggaaccug 7020
agagccuuca cggaggcuau gaccagguau ucugccccuc cugggugaccc ccccagaccg 7080
gaauaugacc uggagcuaau aacaucuugu uccucaaacg ugucuguggc acuuggcccca 7140
cagggccgcc gcagauacua ccugaccaga gaccccacca cuucaauugc ccgggcugcc 7200
uggggaaacag uuagacacuc cccugucaau ucauggcugg gaaacaucau ccaguacgcu 7260
ccaaccauau ggguucgcau gguccugaug acacacuucu ucuccauucu caugggcccag 7320
gacacccuag accagaaccu uaacuuugaa auguacggau cggguguacuc cgugagucccu 7380
cuggaccucc cagccauaau ugaaagguua cacggggcuug acgccuucuc ucugcacaca 7440
uacacucccc acgaacugac gcggguggcu ucagcccuca gaaaacuugg ggcgccaccc 7500
cucagagcgu ggaagagucg ggcgcgugca guuagggcgu cccucaucuc ccguggggggg 7560
agggcggccg uuugcggucg guaccucuuc aacuggggcgg ugaagaccaa gcucaaacuc 7620
acuccuuugc cggaggcacg ccuccuggau uuguccaguu gguuuaccgu cggcgccggc 7680
gggggcgaca uuuaucacag cgugucgcgu gcccgacccc gccauuacu ccuuagccua 7740
cuccuacuuu cuguaggggu aggccucuuc cuacucccccg cucgauagag cggcacacau 7800
uagcuacacu ccauagcuaa cuguuccuuu uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu 7860
uuuuuuuuuu cuuuuuuuuu uuuuucccuc uuucuucccu ucucaucuua uucuacuuuc 7920
uuucuuggug gcuccaucuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980
augacugcag agagugccgu aacuggucuc ucugcagauc augu 8024
```

<210> 3
<211> 9678
<212> DNA
<213> Hepatitis C virus

<220>
<221> CDS
<222> (341)..(9442)

<400> 3
```
acctgcccct aataggggcg acactccgcc atgaatcact cccctgtgag gaactactgt 60

cttcacgcag aaagcgccta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120

cccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg 180

aagactgggt cctttcttgg ataaacccac tctatgcccg gccatttggg cgtgcccccg 240

caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300

cgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat cct 355
                                          Met Ser Thr Asn Pro
```

6

```
aaa cct caa aga aaa acc aaa aga aac acc aac cgt cgc cca gaa gac    403
Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Glu Asp
              10                  15                  20

gtt aag ttc ccg ggc ggc ggc cag atc gtt ggc gga gta tac ttg ttg    451
Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu
              25                  30                  35

ccg cgc agg ggc ccc agg ttg ggt gtg cgc acg aca agg aaa act tcg    499
Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thr Thr Arg Lys Thr Ser
              40                  45                  50

gag cgg tcc cag cca cgt ggg aga cgc cag ccc atc ccc aaa gat cgg    547
Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Asp Arg
          55                  60                  65

cgc tcc act ggc aag gcc tgg gga aaa cca ggt cgc ccc tgg ccc cta    595
Arg Ser Thr Gly Lys Ala Trp Gly Lys Pro Gly Arg Pro Trp Pro Leu
      70                  75                  80                  85

tat ggg aat gag gga ctc ggc tgg gca gga tgg ctc ctg tcc ccc cga    643
Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg
                  90                  95                  100

ggc tct cgc ccc tcc tgg ggc ccc act gac ccc cgg cat agg tcg cgc    691
Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro Arg His Arg Ser Arg
              105                 110                 115

aac gtg ggt aaa gtc atc gac acc cta acg tgt ggc ttt gcc gac ctc    739
Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
          120                 125                 130

atg ggg tac atc ccc gtc gta ggc gcc ccg ctt agt ggc gcc gcc aga    787
Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu Ser Gly Ala Ala Arg
          135                 140                 145

gct gtc gcg cac ggc gtg aga gtc ctg gag gac ggg gtt aat tat gca    835
Ala Val Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala
150                 155                 160                 165

aca ggg aac cta ccc ggt ttc ccc ttt tct atc ttc ttg ctg gcc ctg    883
Thr Gly Asn Leu Pro Gly Phe Pro Phe Ser Ile Phe Leu Leu Ala Leu
                  170                 175                 180

ttg tcc tgc atc acc gtt ccg gtc tct gct gcc cag gtg aag aat acc    931
Leu Ser Cys Ile Thr Val Pro Val Ser Ala Ala Gln Val Lys Asn Thr
              185                 190                 195

agt agc agc tac atg gtg acc aat gac tgc tcc aat gac agc atc act    979
Ser Ser Ser Tyr Met Val Thr Asn Asp Cys Ser Asn Asp Ser Ile Thr
              200                 205                 210

tgg cag ctc gag gct gcg gtt ctc cac gtc ccc ggg tgc gtc ccg tgc    1027
Trp Gln Leu Glu Ala Ala Val Leu His Val Pro Gly Cys Val Pro Cys
          215                 220                 225
```

```
gag aga gtg ggg aat acg tca cgg tgt tgg gtg cca gtc tcg cca aac    1075
Glu Arg Val Gly Asn Thr Ser Arg Cys Trp Val Pro Val Ser Pro Asn
230             235             240             245

atg gct gtg cgg cag ccc ggt gcc ctc acg cag ggt ctg cgg acg cac    1123
Met Ala Val Arg Gln Pro Gly Ala Leu Thr Gln Gly Leu Arg Thr His
                250             255             260

atc gat atg gtt gtg atg tcc gcc acc ttc tgc tct gct ctc tac gtg    1171
Ile Asp Met Val Val Met Ser Ala Thr Phe Cys Ser Ala Leu Tyr Val
            265             270             275

ggg gac ctc tgt ggc ggg gtg atg ctc gcg gcc cag gtg ttc atc gtc    1219
Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala Gln Val Phe Ile Val
            280             285             290

tcg ccg cag tac cac tgg ttt gtg caa gaa tgc aat tgc tcc atc tac    1267
Ser Pro Gln Tyr His Trp Phe Val Gln Glu Cys Asn Cys Ser Ile Tyr
            295             300             305

cct ggc acc atc act gga cac cgc atg gca tgg gac atg atg atg aac    1315
Pro Gly Thr Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
310             315             320             325

tgg tcg ccc acg gcc acc atg atc ctg gcg tac gtg atg cgc gtc ccc    1363
Trp Ser Pro Thr Ala Thr Met Ile Leu Ala Tyr Val Met Arg Val Pro
                330             335             340

gag gtc atc ata gac atc gtt agc ggg gct cac tgg ggc gtc atg ttc    1411
Glu Val Ile Ile Asp Ile Val Ser Gly Ala His Trp Gly Val Met Phe
            345             350             355

ggc ttg gcc tac ttc tct atg cag gga gcg tgg gcg aag gtc att gtc    1459
Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp Ala Lys Val Ile Val
            360             365             370

atc ctt ctg ctg gcc gct ggg gtg gac gcg ggc acc acc acc gtt gga    1507
Ile Leu Leu Leu Ala Ala Gly Val Asp Ala Gly Thr Thr Thr Val Gly
    375             380             385

ggc gct gtt gca cgt tcc acc aac gtg att gcc ggc gtg ttc agc cat    1555
Gly Ala Val Ala Arg Ser Thr Asn Val Ile Ala Gly Val Phe Ser His
390             395             400             405

ggc cct cag cag aac att cag ctc att aac acc aac ggc agt tgg cac    1603
Gly Pro Gln Gln Asn Ile Gln Leu Ile Asn Thr Asn Gly Ser Trp His
                410             415             420

atc aac cgt act gcc ttg aat tgc aat gac tcc ttg aac acc ggc ttt    1651
Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu Asn Thr Gly Phe
            425             430             435

ctc gcg gcc ttg ttc tac acc aac cgc ttt aac tcg tca ggg tgt cca    1699
Leu Ala Ala Leu Phe Tyr Thr Asn Arg Phe Asn Ser Ser Gly Cys Pro
            440             445             450
```

8

```
ggg cgc ctg tcc gcc tgc cgc aac atc gag gct ttc cgg ata ggg tgg    1747
Gly Arg Leu Ser Ala Cys Arg Asn Ile Glu Ala Phe Arg Ile Gly Trp
    455             460             465

ggc acc cta cag tac gag gat aat gtc acc aat cca gag gat atg agg    1795
Gly Thr Leu Gln Tyr Glu Asp Asn Val Thr Asn Pro Glu Asp Met Arg
470             475             480             485

ccg tac tgc tgg cac tac ccc cca aag ccg tgt ggc gta gtc ccc gcg    1843
Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys Gly Val Val Pro Ala
                490             495             500

agg tct gtg tgt ggc cca gtg tac tgt ttc acc ccc agc ccg gta gta    1891
Arg Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val
            505             510             515

gtg ggc acg acc gac aga cgt gga gtg ccc acc tac aca tgg gga gag    1939
Val Gly Thr Thr Asp Arg Arg Gly Val Pro Thr Tyr Thr Trp Gly Glu
            520             525             530

aat gag aca gat gtc ttc cta ctg aac agc acc cga ccg ccg cag ggc    1987
Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr Arg Pro Pro Gln Gly
            535             540             545

tca tgg ttc ggc tgc acg tgg atg aac tcc act ggt ttc acc aag act    2035
Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys Thr
550             555             560             565

tgt ggc gcg cca cct tgc cgc acc aga gct gac ttc aac gcc agc acg    2083
Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp Phe Asn Ala Ser Thr
                570             575             580

gac ttg ttg tgc cct acg gat tgt ttt agg aag cat cct gat gcc act    2131
Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Asp Ala Thr
                585             590             595

tat att aag tgt ggt tct ggg ccc tgg ctc aca cca aag tgc ctg gtc    2179
Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Lys Cys Leu Val
            600             605             610

cac tac cct tac aga ctc tgg cat tac ccc tgc aca gtc aat ttt acc    2227
His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Phe Thr
            615             620             625

atc ttc aag ata aga atg tat gta ggg ggg gtt gag cac agg ctc acg    2275
Ile Phe Lys Ile Arg Met Tyr Val Gly Gly Val Glu His Arg Leu Thr
630             635             640             645

gcc gca tgc aac ttc act cgt ggg gat cgc tgc gac ttg gag gac agg    2323
Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys Asp Leu Glu Asp Arg
                650             655             660

gac agg agt cag ctg tct cct ctg ttg cac tct acc acg gaa tgg gcc    2371
Asp Arg Ser Gln Leu Ser Pro Leu Leu His Ser Thr Thr Glu Trp Ala
            665             670             675

atc ctg ccc tgc acc tac tca gac tta ccc gct ttg tca act ggt ctt    2419
```

9

```
Ile Leu Pro Cys Thr Tyr Ser Asp Leu Pro Ala Leu Ser Thr Gly Leu
        680                 685                 690

ctc cac ctt cac cag aac atc gtg gac gta caa tac atg tat ggc ctc    2467
Leu His Leu His Gln Asn Ile Val Asp Val Gln Tyr Met Tyr Gly Leu
        695                 700                 705

tca cct gct atc aca aaa tac gtc gtt cga tgg gag tgg gtg gta ctc    2515
Ser Pro Ala Ile Thr Lys Tyr Val Val Arg Trp Glu Trp Val Val Leu
710                 715                 720                 725

tta ttc ctg ctc tta gcg gac gcc aga gtc tgc gcc tgc ttg tgg atg    2563
Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp Met
                730                 735                 740

ctc atc ttg ttg ggc cag gcc gaa gca gca ttg gag aag ttg gtc gtc    2611
Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu Glu Lys Leu Val Val
                745                 750                 755

ttg cac gct gcg agt gcg gct aac tgc cat ggc ctc cta tat ttt gcc    2659
Leu His Ala Ala Ser Ala Ala Asn Cys His Gly Leu Leu Tyr Phe Ala
        760                 765                 770

atc ttc ttc gtg gca gct tgg cac atc agg ggt cgg gtg gtc ccc ttg    2707
Ile Phe Phe Val Ala Ala Trp His Ile Arg Gly Arg Val Val Pro Leu
        775                 780                 785

acc acc tat tgc ctc act ggc cta tgg ccc ttc tgc cta ctg ctc atg    2755
Thr Thr Tyr Cys Leu Thr Gly Leu Trp Pro Phe Cys Leu Leu Leu Met
790                 795                 800                 805

gca ctg ccc cgg cag gct tat gcc tat gac gca cct gtg cac gga cag    2803
Ala Leu Pro Arg Gln Ala Tyr Ala Tyr Asp Ala Pro Val His Gly Gln
                810                 815                 820

ata ggc gtg ggt ttg ttg ata ttg atc acc ctc ttc aca ctc acc ccg    2851
Ile Gly Val Gly Leu Leu Ile Leu Ile Thr Leu Phe Thr Leu Thr Pro
                825                 830                 835

ggg tat aag acc ctc ctc ggc cag tgt ctg tgg tgg ttg tgc tat ctc    2899
Gly Tyr Lys Thr Leu Leu Gly Gln Cys Leu Trp Trp Leu Cys Tyr Leu
        840                 845                 850

ctg acc ctg ggg gaa gcc atg att cag gag tgg gta cca ccc atg cag    2947
Leu Thr Leu Gly Glu Ala Met Ile Gln Glu Trp Val Pro Pro Met Gln
        855                 860                 865

gtg cgc ggc ggc cgc gat ggc atc gcg tgg gcc gtc act ata ttc tgc    2995
Val Arg Gly Gly Arg Asp Gly Ile Ala Trp Ala Val Thr Ile Phe Cys
870                 875                 880                 885

ccg ggt gtg gtg ttt gac att acc aaa tgg ctt ttg gcg ttg ctt ggg    3043
Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu Leu Ala Leu Leu Gly
                890                 895                 900

cct gct tac ctc tta agg gcc gct ttg aca cat gtg ccg tac ttc gtc    3091
Pro Ala Tyr Leu Leu Arg Ala Ala Leu Thr His Val Pro Tyr Phe Val
```

10

```
aga gct cac gct ctg ata agg gta tgc gct ttg gtg aag cag ctc gcg    3139
Arg Ala His Ala Leu Ile Arg Val Cys Ala Leu Val Lys Gln Leu Ala
        920                    925                    930

ggg ggt agg tat gtt cag gtg gcg cta ttg gcc ctt ggc agg tgg act    3187
Gly Gly Arg Tyr Val Gln Val Ala Leu Leu Ala Leu Gly Arg Trp Thr
        935                    940                    945

ggc acc tac atc tat gac cac ctc aca cct atg tcg gac tgg gcc gct    3235
Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met Ser Asp Trp Ala Ala
950                    955                    960                    965

agc ggc ctg cgc gac tta gcg gtc gcc gtg gaa ccc atc atc ttc agt    3283
Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Ile Ile Phe Ser
                970                    975                    980

ccg atg gag aag aag gtc atc gtc tgg gga gcg gag acg gct gca tgt    3331
Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala Glu Thr Ala Ala Cys
            985                    990                    995

ggg gac att cta cat gga ctt ccc gtg tcc gcc cga ctc ggc cag gag    3379
Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala Arg Leu Gly Gln Glu
        1000                   1005                   1010

atc ctc ctc ggc cca gct gat ggc tac acc tcc aag ggg tgg aag ctc    3427
Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser Lys Gly Trp Lys Leu
        1015                   1020                   1025

ctt gct ccc atc act gct tat gcc cag caa aca cga ggc ctc ctg ggc    3475
Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly
1030                   1035                   1040                   1045

gcc ata gtg gtg agt atg acg ggg cgt gac agg aca gaa cag gcc ggg    3523
Ala Ile Val Val Ser Met Thr Gly Arg Asp Arg Thr Glu Gln Ala Gly
                1050                   1055                   1060

gaa gtc caa atc ctg tcc aca gtc tct cag tcc ttc ctc gga aca acc    3571
Glu Val Gln Ile Leu Ser Thr Val Ser Gln Ser Phe Leu Gly Thr Thr
                1065                   1070                   1075

atc tcg ggg gtt ttg tgg act gtt tac cac gga gct ggc aac aag act    3619
Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly Ala Gly Asn Lys Thr
            1080                   1085                   1090

cta gcc ggc tta cgg ggt ccg gtc acg cag atg tac tcg agt gct gag    3667
Leu Ala Gly Leu Arg Gly Pro Val Thr Gln Met Tyr Ser Ser Ala Glu
        1095                   1100                   1105

ggg gac ttg gta ggc tgg ccc agc ccc cct ggg acc aag tct ttg gag    3715
Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly Thr Lys Ser Leu Glu
1110                   1115                   1120                   1125

ccg tgc aag tgt gga gcc gtc gac cta tat ctg gtc acg cgg aac gct    3763
Pro Cys Lys Cys Gly Ala Val Asp Leu Tyr Leu Val Thr Arg Asn Ala
            1130                   1135                   1140
```

```
gat gtc atc ccg gct cgg aga cgc ggg gac aag cgg gga gca ttg ctc     3811
Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys Arg Gly Ala Leu Leu
            1145            1150            1155

tcc ccg aga ccc att tcg acc ttg aag ggg tcc tcg ggg ggg ccg gtg     3859
Ser Pro Arg Pro Ile Ser Thr Leu Lys Gly Ser Ser Gly Gly Pro Val
            1160            1165            1170

ctc tgc cct agg ggc cac gtc gtt ggg ctc ttc cga gca gct gtg tgc     3907
Leu Cys Pro Arg Gly His Val Val Gly Leu Phe Arg Ala Ala Val Cys
        1175            1180            1185

tct cgg ggc gtg gcc aaa tcc atc gat ttc atc ccc gtt gag aca ctc     3955
Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile Pro Val Glu Thr Leu
1190            1195            1200            1205

gac gtt gtt aca agg tct ccc act ttc agt gac aac agc acg cca ccg     4003
Asp Val Val Thr Arg Ser Pro Thr Phe Ser Asp Asn Ser Thr Pro Pro
            1210            1215            1220

gct gtg ccc cag acc tat cag gtc ggg tac ttg cat gct cca act ggc     4051
Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu His Ala Pro Thr Gly
            1225            1230            1235

agt gga aag agc acc aag gtc cct gtc gcg tat gcc gcc cag ggg tac     4099
Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr Ala Ala Gln Gly Tyr
            1240            1245            1250

aaa gta cta gtg ctt aac ccc tcg gta gct gcc acc ctg ggg ttt ggg     4147
Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly
        1255            1260            1265

gcg tac cta tcc aag gca cat ggc atc aat ccc aac att agg act gga     4195
Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro Asn Ile Arg Thr Gly
1270            1275            1280            1285

gtc agg acc gtg atg acc ggg gag gcc atc acg tac tcc aca tat ggc     4243
Val Arg Thr Val Met Thr Gly Glu Ala Ile Thr Tyr Ser Thr Tyr Gly
            1290            1295            1300

aaa ttt ctc gcc gat ggg ggc tgc gct agc ggc gcc tat gac atc atc     4291
Lys Phe Leu Ala Asp Gly Gly Cys Ala Ser Gly Ala Tyr Asp Ile Ile
            1305            1310            1315

ata tgc gat gaa tgc cac gct gtg gat gct acc tcc att ctc ggc atc     4339
Ile Cys Asp Glu Cys His Ala Val Asp Ala Thr Ser Ile Leu Gly Ile
            1320            1325            1330

gga acg gtc ctt gat caa gca gag aca gcc ggg gtc aga cta act gtg     4387
Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Val Arg Leu Thr Val
        1335            1340            1345

ctg gct acg gcc aca ccc ccc ggg tca gtg aca acc ccc cat ccc gat     4435
Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Thr Pro His Pro Asp
1350            1355            1360            1365
```

12

```
ata gaa gag gta ggc ctc ggg cgg gag ggt gag atc ccc ttc tat ggg    4483
Ile Glu Glu Val Gly Leu Gly Arg Glu Gly Glu Ile Pro Phe Tyr Gly
              1370              1375              1380

agg gcg att ccc cta tcc tgc atc aag gga ggg aga cac ctg att ttc    4531
Arg Ala Ile Pro Leu Ser Cys Ile Lys Gly Gly Arg His Leu Ile Phe
              1385              1390              1395

tgc cac tca aag aaa aag tgt gac gag ctc gcg gcg gcc ctt cgg ggc    4579
Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Ala Leu Arg Gly
           1400              1405              1410

atg ggc ttg aat gcc gtg gca tac tat aga ggg ttg gac gtc tcc ata    4627
Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Ile
        1415              1420              1425

ata cca gct cag gga gat gtg gtg gtc gtc gcc acc gac gcc ctc atg    4675
Ile Pro Ala Gln Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met
1430              1435              1440              1445

acg ggg tac act gga gac ttt gac tcc gtg atc gac tgc aat gta gcg    4723
Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Val Ala
              1450              1455              1460

gtc acc caa gct gtc gac ttc agc ctg gac ccc acc ttc act ata acc    4771
Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Thr
              1465              1470              1475

aca cag act gtc cca caa gac gct gtc tca cgc agt cag cgc cgc ggg    4819
Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly
           1480              1485              1490

cgc aca ggt aga gga aga cag ggc act tat agg tat gtt tcc act ggt    4867
Arg Thr Gly Arg Gly Arg Gln Gly Thr Tyr Arg Tyr Val Ser Thr Gly
        1495              1500              1505

gaa cga gcc tca gga atg ttt gac agt gta gtg ctt tgt gag tgc tac    4915
Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val Leu Cys Glu Cys Tyr
1510              1515              1520              1525

gac gca ggg gct gcg tgg tac gat ctc aca cca gcg gag acc acc gtc    4963
Asp Ala Gly Ala Ala Trp Tyr Asp Leu Thr Pro Ala Glu Thr Thr Val
           1530              1535              1540

agg ctt aga gcg tat ttc aac acg ccc ggc cta ccc gtg tgt caa gac    5011
Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu Pro Val Cys Gln Asp
              1545              1550              1555

cat ctt gaa ttt tgg gag gca gtt ttc acc ggc ctc aca cac ata gac    5059
His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly Leu Thr His Ile Asp
           1560              1565              1570

gcc cac ttc ctc tcc caa aca aag caa gcg ggg gag aac ttc gcg tac    5107
Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly Glu Asn Phe Ala Tyr
        1575              1580              1585

cta gta gcc tac caa gct acg gtg tgc gcc aga gcc aag gcc cct ccc    5155
```

13

```
Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Lys Ala Pro Pro
1590             1595             1600             1605


ccg tcc tgg gac gcc atg tgg aag tgc ctg gcc cga ctc aag cct acg   5203
Pro Ser Trp Asp Ala Met Trp Lys Cys Leu Ala Arg Leu Lys Pro Thr
            1610             1615             1620


ctt gcg ggc ccc aca cct ctc ctg tac cgt ttg ggc cct att acc aat   5251
Leu Ala Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Pro Ile Thr Asn
            1625             1630             1635


gag gtc acc ctc aca cac cct ggg acg aag tac atc gcc aca tgc atg   5299
Glu Val Thr Leu Thr His Pro Gly Thr Lys Tyr Ile Ala Thr Cys Met
        1640             1645             1650


caa gct gac ctt gag gtc atg acc agc acg tgg gtc cta gct gga gga   5347
Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp Val Leu Ala Gly Gly
    1655             1660             1665


gtc ctg gca gcc gtc gcc gca tat tgc ctg gcg act gga tgc gtt tcc   5395
Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala Thr Gly Cys Val Ser
1670             1675             1680             1685


atc atc ggc cgc ttg cac gtc aac cag cga gtc gtc gtt gcg ccg gat   5443
Ile Ile Gly Arg Leu His Val Asn Gln Arg Val Val Val Ala Pro Asp
            1690             1695             1700


aag gag gtc ctg tat gag gct ttt gat gag atg gag gaa tgc gcc tct   5491
Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met Glu Glu Cys Ala Ser
            1705             1710             1715


agg gcg gct ctc atc gaa gag ggg cag cgg ata gcc gag atg ttg aag   5539
Arg Ala Ala Leu Ile Glu Glu Gly Gln Arg Ile Ala Glu Met Leu Lys
            1720             1725             1730


tcc aag atc caa ggc ttg ctg cag cag gcc tct aag cag gcc cag gac   5587
Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser Lys Gln Ala Gln Asp
    1735             1740             1745


ata caa ccc gct atg cag gct tca tgg ccc aaa gtg gaa caa ttt tgg   5635
Ile Gln Pro Ala Met Gln Ala Ser Trp Pro Lys Val Glu Gln Phe Trp
1750             1755             1760             1765


gcc aga cac atg tgg aac ttc att agc ggc atc caa tac ctc gca gga   5683
Ala Arg His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly
            1770             1775             1780


ttg tca aca ctg cca ggg aac ccc gcg gtg gct tcc atg atg gca ttc   5731
Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala Ser Met Met Ala Phe
            1785             1790             1795


agt gcc gcc ctc acc agt ccg ttg tcg acc agt acc acc atc ctt ctc   5779
Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser Thr Thr Ile Leu Leu
        1800             1805             1810


aac atc atg gga ggc tgg tta gcg tcc cag atc gca cca ccc gcg ggg   5827
Asn Ile Met Gly Gly Trp Leu Ala Ser Gln Ile Ala Pro Pro Ala Gly
```

14

```
gcc acc ggc ttt gtc gtc agt ggc ctg gtg ggg gct gcc gtg ggc agc     5875
Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly Ala Ala Val Gly Ser
1830                     1835                     1840                     1845
```

```
ata ggc ctg ggt aag gtg ctg gtg gac atc ctg gca gga tat ggt gcg     5923
Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly Ala
                1850                     1855                     1860
```

```
ggc att tcg ggg gcc ctc gtc gca ttc aag atc atg tct ggc gag aag     5971
Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Lys
                1865                     1870                     1875
```

```
ccc tct atg gaa gat gtc atc aat cta ctg cct ggg atc ctg tct ccg     6019
Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro Gly Ile Leu Ser Pro
            1880                     1885                     1890
```

```
gga gcc ctg gtg gtg ggg gtc atc tgc gcg gcc att ctg cgc cgc cac     6067
Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala Ile Leu Arg Arg His
        1895                     1900                     1905
```

```
gtg gga ccg ggg gag ggc gcg gtc caa tgg atg aac agg ctt att gcc     6115
Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala
1910                     1915                     1920                     1925
```

```
ttt gct tcc aga gga aac cac gtc gcc cct act cac tac gtg acg gag     6163
Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr His Tyr Val Thr Glu
                1930                     1935                     1940
```

```
tcg gat gcg tcg cag cgt gtg acc caa cta ctt ggc tct ctt act ata     6211
Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu Gly Ser Leu Thr Ile
            1945                     1950                     1955
```

```
acc agc cta ctc aga aga ctc cac aat tgg ata act gag gac tgc ccc     6259
Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile Thr Glu Asp Cys Pro
        1960                     1965                     1970
```

```
atc cca tgc tcc gga tcc tgg ctc cgc gac gtg tgg gac tgg gtt tgc     6307
Ile Pro Cys Ser Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Val Cys
```

```
acc atc ttg aca gac ttc aaa aat tgg ctg acc tct aaa ttg ttc ccc     6355
Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr Ser Lys Leu Phe Pro
1990                     1995                     2000                     2005
```

```
aag ctg ccc ggc ctc ccc ttc atc tct tgt caa aag ggg tac aag ggt     6403
Lys Leu Pro Gly Leu Pro Phe Ile Ser Cys Gln Lys Gly Tyr Lys Gly
                2010                     2015                     2020
```

```
gtg tgg gcc ggc act ggc atc atg acc acg cgc tgc cct tgc ggc gcc     6451
Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg Cys Pro Cys Gly Ala
                2025                     2030                     2035
```

```
aac atc tct ggc aat gtc cgc ctg ggc tct atg agg atc aca ggg cct     6499
Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met Arg Ile Thr Gly Pro
```

15

```
aaa acc tgc atg aac acc tgg cag ggg acc ttt cct atc aat tgc tac   6547
Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe Pro Ile Asn Cys Tyr
  2055                2060                2065
```

```
acg gag ggc cag tgc gcg ccg aaa ccc ccc acg aac tac aag acc gcc   6595
Thr Glu Gly Gln Cys Ala Pro Lys Pro Pro Thr Asn Tyr Lys Thr Ala
2070                2075                2080                2085
```

```
atc tgg agg gtg gcg gcc tcg gag tac gcg gag gtg acg cag cat ggg   6643
Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu Val Thr Gln His Gly
              2090                2095                2100
```

```
tcg tac tcc tat gta aca gga ctg acc act gac aat ctg aaa att cct   6691
Ser Tyr Ser Tyr Val Thr Gly Leu Thr Thr Asp Asn Leu Lys Ile Pro
              2105                2110                2115
```

```
tgc caa cta cct tct cca gag ttt ttc tcc tgg gtg gac ggt gtg cag   6739
Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp Val Asp Gly Val Gln
          2120                2125                2130
```

```
atc cat agg ttt gca ccc aca cca aag ccg ttt ttc cgg gat gag gtc   6787
Ile His Arg Phe Ala Pro Thr Pro Lys Pro Phe Phe Arg Asp Glu Val
      2135                2140                2145
```

```
tcg ttc tgc gtt ggg ctt aat tcc tat gct gtc ggg tcc cag ctt ccc   6835
Ser Phe Cys Val Gly Leu Asn Ser Tyr Ala Val Gly Ser Gln Leu Pro
2150                2155                2160                2165
```

```
tgt gaa cct gag ccc gac gca gac gta ttg agg tcc atg cta aca gat   6883
Cys Glu Pro Glu Pro Asp Ala Asp Val Leu Arg Ser Met Leu Thr Asp
              2170                2175                2180
```

```
ccg ccc cac atc acg gcg gag act gcg gcg cgg cgc ttg gca cgg gga   6931
Pro Pro His Ile Thr Ala Glu Thr Ala Ala Arg Arg Leu Ala Arg Gly
              2185                2190                2195
```

```
tca cct cca tct gag gcg agc tcc tca gtg agc cag cta tca gca ccg   6979
Ser Pro Pro Ser Glu Ala Ser Ser Ser Val Ser Gln Leu Ser Ala Pro
          2200                2205                2210
```

```
tcg ctg cgg gcc acc tgc acc acc cac agc aac acc tat gac gtg gac   7027
Ser Leu Arg Ala Thr Cys Thr Thr His Ser Asn Thr Tyr Asp Val Asp
      2215                2220                2225
```

```
atg gtc gat gcc aac ctg ctc atg gag ggc ggt gtg gct cag aca gag   7075
Met Val Asp Ala Asn Leu Leu Met Glu Gly Gly Val Ala Gln Thr Glu
2230                2235                2240                2245
```

```
cct gag tcc agg gtg ccc gtt ctg gac ttt ctc gag cca atg gcc gag   7123
Pro Glu Ser Arg Val Pro Val Leu Asp Phe Leu Glu Pro Met Ala Glu
              2250                2255                2260
```

```
gaa gag agc gac ctt gag ccc tca ata cca tcg gag tgc atg ctc ccc   7171
Glu Glu Ser Asp Leu Glu Pro Ser Ile Pro Ser Glu Cys Met Leu Pro
          2265                2270                2275
```

```
agg agc ggg ttt cca cgg gcc tta ccg gct tgg gca cgg cct gac tac     7219
Arg Ser Gly Phe Pro Arg Ala Leu Pro Ala Trp Ala Arg Pro Asp Tyr
        2280                2285                2290


aac ccg ccg ctc gtg gaa tcg tgg agg agg cca gat tac caa ccg ccc     7267
Asn Pro Pro Leu Val Glu Ser Trp Arg Arg Pro Asp Tyr Gln Pro Pro
        2295                2300                2305


acc gtt gct ggt tgt gct ctc ccc ccc ccc aag aag gcc ccg acg cct     7315
Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys Lys Ala Pro Thr Pro
2310                2315                2320                2325


ccc cca agg aga cgc cgg aca gtg ggt ctg agc gag agc acc ata tca     7363
Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser Glu Ser Thr Ile Ser
            2330                2335                2340


gaa gcc ctc cag caa ctg gcc atc aag acc ttt ggc cag ccc ccc tcg     7411
Glu Ala Leu Gln Gln Leu Ala Ile Lys Thr Phe Gly Gln Pro Pro Ser
            2345                2350                2355


agc ggt gat gca ggc tcg tcc acg ggg gcg ggc gcc gcc gaa tcc ggc     7459
Ser Gly Asp Ala Gly Ser Ser Thr Gly Ala Gly Ala Ala Glu Ser Gly
        2360                2365                2370


ggt ccg acg tcc cct ggt gag ccg gcc ccc tca gag aca ggt tcc gcc     7507
Gly Pro Thr Ser Pro Gly Glu Pro Ala Pro Ser Glu Thr Gly Ser Ala
    2375                2380                2385


tcc tct atg ccc ccc ctc gag ggg gag cct gga gat ccg gac ctg gag     7555
Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Glu
2390                2395                2400                2405


tct gat cag gta gag ctt caa cct ccc ccc cag ggg ggg ggg gta gct     7603
Ser Asp Gln Val Glu Leu Gln Pro Pro Pro Gln Gly Gly Gly Val Ala
            2410                2415                2420


ccc ggt tcg ggc tcg ggg tct tgg tct act tgc tcc gag gag gac gat     7651
Pro Gly Ser Gly Ser Gly Ser Trp Ser Thr Cys Ser Glu Glu Asp Asp
            2425                2430                2435


acc acc gtg tgc tgc tcc atg tca tac tcc tgg acc ggg gct cta ata     7699
Thr Thr Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Ile
        2440                2445                2450


act ccc tgt agc ccc gaa gag gaa aag ttg cca atc aac cct ttg agt     7747
Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro Ile Asn Pro Leu Ser
    2455                2460                2465


aac tcg ctg ttg cga tac cat aac aag gtg tac tgt aca aca tca aag     7795
Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr Cys Thr Thr Ser Lys
2470                2475                2480                2485


agc gcc tca cag agg gct aaa aag gta act ttt gac agg acg caa gtg     7843
Ser Ala Ser Gln Arg Ala Lys Lys Val Thr Phe Asp Arg Thr Gln Val
            2490                2495                2500
```

17

```
ctc gac gcc cat tat gac tca gtc tta aag gac atc aag cta gcg gct    7891
Leu Asp Ala His Tyr Asp Ser Val Leu Lys Asp Ile Lys Leu Ala Ala
            2505              2510              2515

tcc aag gtc agc gca agg ctc ctc acc ttg gag gag gcg tgc cag ttg    7939
Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu Glu Ala Cys Gln Leu
            2520              2525              2530

act cca ccc cat tct gca aga tcc aag tat gga ttc ggg gcc aag gag    7987
Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly Phe Gly Ala Lys Glu
          2535              2540              2545

gtc cgc agc ttg tcc ggg agg gcc gtt aac cac atc aag tcc gtg tgg    8035
Val Arg Ser Leu Ser Gly Arg Ala Val Asn His Ile Lys Ser Val Trp
2550              2555              2560              2565

aag gac ctc ctg gaa gac cca caa aca cca att ccc aca acc atc atg    8083
Lys Asp Leu Leu Glu Asp Pro Gln Thr Pro Ile Pro Thr Thr Ile Met
                2570              2575              2580

gcc aaa aat gag gtg ttc tgc gtg gac ccc gcc aag ggg ggt aag aaa    8131
Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala Lys Gly Gly Lys Lys
              2585              2590              2595

cca gct cgc ctc atc gtt tac cct gac ctc ggc gtc cgg gtc tgc gag    8179
Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu
          2600              2605              2610

aaa atg gcc ctc tat gac att aca caa aag ctt cct cag gcg gta atg    8227
Lys Met Ala Leu Tyr Asp Ile Thr Gln Lys Leu Pro Gln Ala Val Met
    2615              2620              2625

gga gct tcc tat ggc ttc cag tac tcc cct gcc caa cgg gtg gag tat    8275
Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Glu Tyr
2630              2635              2640              2645

ctc ttg aaa gca tgg gcg gaa aag aag gac ccc atg ggt ttt tcg tat    8323
Leu Leu Lys Ala Trp Ala Glu Lys Lys Asp Pro Met Gly Phe Ser Tyr
              2650              2655              2660

gat acc cga tgc ttc gac tca acc gtc act gag aga gac atc agg acc    8371
Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp Ile Arg Thr
              2665              2670              2675

gag gag tcc ata tac cag gcc tgc tcc ctg ccc gag gag gcc cgc act    8419
Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu Glu Ala Arg Thr
          2680              2685              2690

gcc ata cac tcg ctg act gag aga ctt tac gta gga ggg ccc atg ttc    8467
Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Met Phe
    2695              2700              2705

aac agc aag ggt caa acc tgc ggt tac aga cgt tgc cgc gcc agc ggg    8515
Asn Ser Lys Gly Gln Thr Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly
2710              2715              2720              2725

gtg cta acc act agc atg ggt aac acc atc aca tgc tat gtg aaa gcc    8563
```

18

```
Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr Cys Tyr Val Lys Ala
        2730              2735              2740

cta gcg gcc tgc aag gct gcg ggg ata gtt gcg ccc aca atg ctg gta    8611
Leu Ala Ala Cys Lys Ala Ala Gly Ile Val Ala Pro Thr Met Leu Val
        2745              2750              2755

tgc ggc gat gac cta gta gtc atc tca gaa agc cag ggg act gag gag    8659
Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser Gln Gly Thr Glu Glu
        2760              2765              2770

gac gag cgg aac ctg aga gcc ttc acg gag gcc atg acc agg tac tct    8707
Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser
      2775              2780              2785

gcc cct cct ggt gat ccc ccc aga ccg gaa tat gac ctg gag cta ata    8755
Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr Asp Leu Glu Leu Ile
2790              2795              2800              2805

aca tcc tgt tcc tca aat gtg tct gtg gcg ttg ggc ccg cgg ggc cgc    8803
Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu Gly Pro Arg Gly Arg
              2810              2815              2820

cgc aga tac tac ctg acc aga gac cca acc act cca ctc gcc cgg gct    8851
Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala
              2825              2830              2835

gcc tgg gaa aca gtt aga cac tcc cct atc aat tca tgg ctg gga aac    8899
Ala Trp Glu Thr Val Arg His Ser Pro Ile Asn Ser Trp Leu Gly Asn
        2840              2845              2850

atc atc cag tat gct cca acc ata tgg gtt cgc atg gtc cta atg aca    8947
Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg Met Val Leu Met Thr
    2855              2860              2865

cac ttc ttc tcc att ctc atg gtc caa gac acc ctg gac cag aac ctc    8995
His Phe Phe Ser Ile Leu Met Val Gln Asp Thr Leu Asp Gln Asn Leu
2870              2875              2880              2885

aac ttt gag atg tat gga tca gta tac tcc gtg aat cct ttg gac ctt    9043
Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Asn Pro Leu Asp Leu
              2890              2895              2900

cca gcc ata att gag agg tta cac ggg ctt gac gcc ttt tct atg cac    9091
Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp Ala Phe Ser Met His
              2905              2910              2915

aca tac tct cac cac gaa ctg acg cgg gtg gct tca gcc ctc aga aaa    9139
Thr Tyr Ser His His Glu Leu Thr Arg Val Ala Ser Ala Leu Arg Lys
        2920              2925              2930

ctt ggg gcg cca ccc ctc agg gtg tgg aag agt cgg gct cgc gca gtc    9187
Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser Arg Ala Arg Ala Val
    2935              2940              2945

agg gcg tcc ctc atc tcc cgt gga ggg aaa gcg gcc gtt tgc ggc cga    9235
Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala Ala Val Cys Gly Arg
```

19

```
tat ctc ttc aat tgg gcg gtg aag acc aag ctc aaa ctc act cca ttg   9283
Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu
            2970                2975                2980

ccg gag gcg cgc cta ctg gac tta tcc agt tgg ttc acc gtc ggc gcc   9331
Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Gly Ala
            2985                2990                2995

ggc ggg ggc gac att ttt cac agc gtg tcg cgc gcc cga ccc cgc tca   9379
Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg Ala Arg Pro Arg Ser
            3000                3005                3010

tta ctc ttc ggc cta ctc cta ctt ttc gta ggg gta ggc ctc ttc cta   9427
Leu Leu Phe Gly Leu Leu Leu Leu Phe Val Gly Val Gly Leu Phe Leu
         3015                3020                3025

ctc ccc gct cgg tag agcggcacac actaggtaca ctccatagct aactgttcct   9482
Leu Pro Ala Arg
3030

tttttttttt tttttttttt tttttttttt tttttttttt ttcttttttt tttttttccc   9542

tctttcttcc cttctcatct tattctactt tctttcttgg tggctccatc ttagccctag   9602

tcacggctag ctgtgaaagg tccgtgagcc gcatgactgc agagagtgcc gtaactggtc   9662

tctctgcaga tcatgt                                                    9678
```

<210> 4
<211> 3033
<212> PRT
<213> Hepatitis C virus

<400> 4
```
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
1               5                   10                  15
Arg Arg Pro Glu Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
            20                  25                  30
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thr
            35                  40                  45
Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
    50                  55                  60
Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ala Trp Gly Lys Pro Gly
65                  70                  75                  80
Arg Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
                85                  90                  95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro
            100                 105                 110
Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
        115                 120                 125
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu
    130                 135                 140
Ser Gly Ala Ala Arg Ala Val Ala His Gly Val Arg Val Leu Glu Asp
```

```
145                      150                      155                      160
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Phe Pro Phe Ser Ile
            165                      170                      175
Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Val Pro Val Ser Ala Ala
            180                      185                      190
Gln Val Lys Asn Thr Ser Ser Ser Tyr Met Val Thr Asn Asp Cys Ser
            195                      200                      205
Asn Asp Ser Ile Thr Trp Gln Leu Glu Ala Ala Val Leu His Val Pro
        210                      215                      220
Gly Cys Val Pro Cys Glu Arg Val Gly Asn Thr Ser Arg Cys Trp Val
225                      230                      235                      240
Pro Val Ser Pro Asn Met Ala Val Arg Gln Pro Gly Ala Leu Thr Gln
                245                      250                      255
Gly Leu Arg Thr His Ile Asp Met Val Val Met Ser Ala Thr Phe Cys
            260                      265                      270
Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala
            275                      280                      285
Gln Val Phe Ile Val Ser Pro Gln Tyr His Trp Phe Val Gln Glu Cys
        290                      295                      300
Asn Cys Ser Ile Tyr Pro Gly Thr Ile Thr Gly His Arg Met Ala Trp
305                      310                      315                      320
Asp Met Met Met Asn Trp Ser Pro Thr Ala Thr Met Ile Leu Ala Tyr
                325                      330                      335
Val Met Arg Val Pro Glu Val Ile Ile Asp Ile Val Ser Gly Ala His
            340                      345                      350
Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp
            355                      360                      365
Ala Lys Val Ile Val Ile Leu Leu Leu Ala Ala Gly Val Asp Ala Gly
        370                      375                      380
Thr Thr Thr Val Gly Gly Ala Val Ala Arg Ser Thr Asn Val Ile Ala
385                      390                      395                      400
Gly Val Phe Ser His Gly Pro Gln Gln Asn Ile Gln Leu Ile Asn Thr
                405                      410                      415
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
            420                      425                      430
Leu Asn Thr Gly Phe Leu Ala Ala Leu Phe Tyr Thr Asn Arg Phe Asn
            435                      440                      445
Ser Ser Gly Cys Pro Gly Arg Leu Ser Ala Cys Arg Asn Ile Glu Ala
        450                      455                      460
Phe Arg Ile Gly Trp Gly Thr Leu Gln Tyr Glu Asp Asn Val Thr Asn
465                      470                      475                      480
Pro Glu Asp Met Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys
                485                      490                      495
Gly Val Val Pro Ala Arg Ser Val Cys Gly Pro Val Tyr Cys Phe Thr
            500                      505                      510
Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Arg Gly Val Pro Thr
            515                      520                      525
Tyr Thr Trp Gly Glu Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr
            530                      535                      540
Arg Pro Pro Gln Gly Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr
545                      550                      555                      560
Gly Phe Thr Lys Thr Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp
                565                      570                      575
Phe Asn Ala Ser Thr Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys
            580                      585                      590
His Pro Asp Ala Thr Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr
            595                      600                      605
```

21

```
Pro Lys Cys Leu Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys
    610                 615                 620
Thr Val Asn Phe Thr Ile Phe Lys Ile Arg Met Tyr Val Gly Gly Val
625                 630                 635                 640
Glu His Arg Leu Thr Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys
                645                 650                 655
Asp Leu Glu Asp Arg Asp Arg Ser Gln Leu Ser Pro Leu Leu His Ser
                660                 665                 670
Thr Thr Glu Trp Ala Ile Leu Pro Cys Thr Tyr Ser Asp Leu Pro Ala
                675                 680                 685
Leu Ser Thr Gly Leu Leu His Leu His Gln Asn Ile Val Asp Val Gln
    690                 695                 700
Tyr Met Tyr Gly Leu Ser Pro Ala Ile Thr Lys Tyr Val Val Arg Trp
705                 710                 715                 720
Glu Trp Val Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys
                725                 730                 735
Ala Cys Leu Trp Met Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu
                740                 745                 750
Glu Lys Leu Val Val Leu His Ala Ala Ser Ala Ala Asn Cys His Gly
    755                 760                 765
Leu Leu Tyr Phe Ala Ile Phe Phe Val Ala Ala Trp His Ile Arg Gly
    770                 775                 780
Arg Val Val Pro Leu Thr Thr Tyr Cys Leu Thr Gly Leu Trp Pro Phe
785                 790                 795                 800
Cys Leu Leu Leu Met Ala Leu Pro Arg Gln Ala Tyr Ala Tyr Asp Ala
                805                 810                 815
Pro Val His Gly Gln Ile Gly Val Gly Leu Leu Ile Leu Ile Thr Leu
                820                 825                 830
Phe Thr Leu Thr Pro Gly Tyr Lys Thr Leu Leu Gly Gln Cys Leu Trp
                835                 840                 845
Trp Leu Cys Tyr Leu Leu Thr Leu Gly Glu Ala Met Ile Gln Glu Trp
    850                 855                 860
Val Pro Pro Met Gln Val Arg Gly Gly Arg Asp Gly Ile Ala Trp Ala
865                 870                 875                 880
Val Thr Ile Phe Cys Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu
                885                 890                 895
Leu Ala Leu Leu Gly Pro Ala Tyr Leu Leu Arg Ala Ala Leu Thr His
    900                 905                 910
Val Pro Tyr Phe Val Arg Ala His Ala Leu Ile Arg Val Cys Ala Leu
    915                 920                 925
Val Lys Gln Leu Ala Gly Gly Arg Tyr Val Gln Val Ala Leu Leu Ala
    930                 935                 940
Leu Gly Arg Trp Thr Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met
945                 950                 955                 960
Ser Asp Trp Ala Ala Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu
                965                 970                 975
Pro Ile Ile Phe Ser Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala
                980                 985                 990
Glu Thr Ala Ala Cys Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala
    995                 1000                1005
Arg Leu Gly Gln Glu Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser
    1010                1015                1020
Lys Gly Trp Lys Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr
1025                1030                1035                1040
Arg Gly Leu Leu Gly Ala Ile Val Val Ser Met Thr Gly Arg Asp Arg
                1045                1050                1055
Thr Glu Gln Ala Gly Glu Val Gln Ile Leu Ser Thr Val Ser Gln Ser
```

22

```
           1060                    1065                    1070
Phe Leu Gly Thr Thr Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly
       1075                    1080                    1085
Ala Gly Asn Lys Thr Leu Ala Gly Leu Arg Gly Pro Val Thr Gln Met
    1090                    1095                    1100
Tyr Ser Ser Ala Glu Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly
1105                    1110                    1115                    1120
Thr Lys Ser Leu Glu Pro Cys Lys Cys Gly Ala Val Asp Leu Tyr Leu
               1125                    1130                    1135
Val Thr Arg Asn Ala Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys
       1140                    1145                    1150
Arg Gly Ala Leu Leu Ser Pro Arg Pro Ile Ser Thr Leu Lys Gly Ser
       1155                    1160                    1165
Ser Gly Gly Pro Val Leu Cys Pro Arg Gly His Val Val Gly Leu Phe
    1170                    1175                    1180
Arg Ala Ala Val Cys Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile
1185                    1190                    1195                    1200
Pro Val Glu Thr Leu Asp Val Val Thr Arg Ser Pro Thr Phe Ser Asp
               1205                    1210                    1215
Asn Ser Thr Pro Pro Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu
               1220                    1225                    1230
His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr
       1235                    1240                    1245
Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala
    1250                    1255                    1260
Thr Leu Gly Phe Gly Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro
1265                    1270                    1275                    1280
Asn Ile Arg Thr Gly Val Arg Thr Val Met Thr Gly Glu Ala Ile Thr
               1285                    1290                    1295
Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ala Ser Gly
           1300                    1305                    1310
Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ala Val Asp Ala Thr
       1315                    1320                    1325
Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly
    1330                    1335                    1340
Val Arg Leu Thr Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr
1345                    1350                    1355                    1360
Thr Pro His Pro Asp Ile Glu Glu Val Gly Leu Gly Arg Glu Gly Glu
               1365                    1370                    1375
Ile Pro Phe Tyr Gly Arg Ala Ile Pro Leu Ser Cys Ile Lys Gly Gly
           1380                    1385                    1390
Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala
       1395                    1400                    1405
Ala Ala Leu Arg Gly Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly
    1410                    1415                    1420
Leu Asp Val Ser Ile Ile Pro Ala Gln Gly Asp Val Val Val Val Ala
1425                    1430                    1435                    1440
Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile
               1445                    1450                    1455
Asp Cys Asn Val Ala Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro
           1460                    1465                    1470
Thr Phe Thr Ile Thr Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg
       1475                    1480                    1485
Ser Gln Arg Arg Gly Arg Thr Gly Arg Gly Arg Gln Gly Thr Tyr Arg
    1490                    1495                    1500
Tyr Val Ser Thr Gly Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val
1505                    1510                    1515                    1520
```

Leu Cys Glu Cys Tyr Asp Ala Gly Ala Ala Trp Tyr Asp Leu Thr Pro
          1525   1530   1535
Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu
          1540   1545   1550
Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly
          1555   1560   1565
Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ala Gly
       1570   1575   1580
Glu Asn Phe Ala Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg
    1585   1590   1595   1600
Ala Lys Ala Pro Pro Pro Ser Trp Asp Ala Met Trp Lys Cys Leu Ala
          1605   1610   1615
Arg Leu Lys Pro Thr Leu Ala Gly Pro Thr Pro Leu Leu Tyr Arg Leu
          1620   1625   1630
Gly Pro Ile Thr Asn Glu Val Thr Leu Thr His Pro Gly Thr Lys Tyr
       1635   1640   1645
Ile Ala Thr Cys Met Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp
    1650   1655   1660
Val Leu Ala Gly Gly Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala
1665   1670   1675   1680
Thr Gly Cys Val Ser Ile Ile Gly Arg Leu His Val Asn Gln Arg Val
          1685   1690   1695
Val Val Ala Pro Asp Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met
          1700   1705   1710
Glu Glu Cys Ala Ser Arg Ala Ala Leu Ile Glu Glu Gly Gln Arg Ile
       1715   1720   1725
Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser
    1730   1735   1740
Lys Gln Ala Gln Asp Ile Gln Pro Ala Met Gln Ala Ser Trp Pro Lys
1745   1750   1755   1760
Val Glu Gln Phe Trp Ala Arg His Met Trp Asn Phe Ile Ser Gly Ile
          1765   1770   1775
Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala
       1780   1785   1790
Ser Met Met Ala Phe Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser
       1795   1800   1805
Thr Thr Ile Leu Leu Asn Ile Met Gly Gly Trp Leu Ala Ser Gln Ile
    1810   1815   1820
Ala Pro Pro Ala Gly Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly
1825   1830   1835   1840
Ala Ala Val Gly Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu
          1845   1850   1855
Ala Gly Tyr Gly Ala Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile
       1860   1865   1870
Met Ser Gly Glu Lys Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro
       1875   1880   1885
Gly Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala
    1890   1895   1900
Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met
1905   1910   1915   1920
Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr
          1925   1930   1935
His Tyr Val Thr Glu Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu
       1940   1945   1950
Gly Ser Leu Thr Ile Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile
       1955   1960   1965
Thr Glu Asp Cys Pro Ile Pro Cys Ser Gly Ser Trp Leu Arg Asp Val

24

```
        1970                1975                1980
Trp Asp Trp Val Cys Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr
1985                1990                1995                2000
Ser Lys Leu Phe Pro Lys Leu Pro Gly Leu Pro Phe Ile Ser Cys Gln
              2005                2010                2015
Lys Gly Tyr Lys Gly Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg
            2020                2025                2030
Cys Pro Cys Gly Ala Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met
          2035                2040                2045
Arg Ile Thr Gly Pro Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe
        2050                2055                2060
Pro Ile Asn Cys Tyr Thr Glu Gly Gln Cys Ala Pro Lys Pro Pro Thr
2065                2070                2075                2080
Asn Tyr Lys Thr Ala Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu
              2085                2090                2095
Val Thr Gln His Gly Ser Tyr Ser Tyr Val Thr Gly Leu Thr Thr Asp
            2100                2105             .   2110
Asn Leu Lys Ile Pro Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp
          2115                2120                2125
Val Asp Gly Val Gln Ile His Arg Phe Ala Pro Thr Pro Lys Pro Phe
        2130                2135                2140


Phe Arg Asp Glu Val Ser Phe Cys Val Gly Leu Asn Ser Tyr Ala Val
2145                2150                2155                2160
Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Ala Asp Val Leu Arg
              2165                2170                2175
Ser Met Leu Thr Asp Pro Pro His Ile Thr Ala Glu Thr Ala Ala Arg
            2180                2185                2190
Arg Leu Ala Arg Gly Ser Pro Pro Ser Glu Ala Ser Ser Ser Val Ser
          2195                2200                2205
Gln Leu Ser Ala Pro Ser Leu Arg Ala Thr Cys Thr Thr His Ser Asn
      2210                2215                2220
Thr Tyr Asp Val Asp Met Val Asp Ala Asn Leu Leu Met Glu Gly Gly
2225                2230                2235                2240
Val Ala Gln Thr Glu Pro Glu Ser Arg Val Pro Val Leu Asp Phe Leu
              2245                2250                2255
Glu Pro Met Ala Glu Glu Glu Ser Asp Leu Glu Pro Ser Ile Pro Ser
            2260                2265                2270
Glu Cys Met Leu Pro Arg Ser Gly Phe Pro Arg Ala Leu Pro Ala Trp
          2275                2280                2285
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Ser Trp Arg Arg Pro
    2290                2295                2300
Asp Tyr Gln Pro Pro Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys
2305                2310                2315                2320
Lys Ala Pro Thr Pro Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser
              2325                2330                2335
Glu Ser Thr Ile Ser Glu Ala Leu Gln Gln Leu Ala Ile Lys Thr Phe
            2340                2345                2350
Gly Gln Pro Pro Ser Ser Gly Asp Ala Gly Ser Ser Thr Gly Ala Gly
          2355                2360                2365
Ala Ala Glu Ser Gly Gly Pro Thr Ser Pro Gly Glu Pro Ala Pro Ser
    2370                2375                2380
Glu Thr Gly Ser Ala Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly
2385                2390                2395                2400
Asp Pro Asp Leu Glu Ser Asp Gln Val Glu Leu Gln Pro Pro Pro Gln
              2405                2410                2415
Gly Gly Gly Val Ala Pro Gly Ser Gly Ser Gly Ser Trp Ser Thr Cys
```

25

```
              2420              2425              2430
Ser Glu Glu Asp Asp Thr Thr Val Cys Cys Ser Met Ser Tyr Ser Trp
       2435              2440              2445
Thr Gly Ala Leu Ile Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro
      2450              2455              2460
Ile Asn Pro Leu Ser Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr
2465              2470              2475              2480
Cys Thr Thr Ser Lys Ser Ala Ser Gln Arg Ala Lys Lys Val Thr Phe
             2485              2490              2495
Asp Arg Thr Gln Val Leu Asp Ala His Tyr Asp Ser Val Leu Lys Asp
         2500              2505              2510
Ile Lys Leu Ala Ala Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu
         2515              2520              2525
Glu Ala Cys Gln Leu Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly
    2530              2535              2540
Phe Gly Ala Lys Glu Val Arg Ser Leu Ser Gly Arg Ala Val Asn His
2545              2550              2555              2560
Ile Lys Ser Val Trp Lys Asp Leu Leu Glu Asp Pro Gln Thr Pro Ile
             2565              2570              2575
Pro Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala
         2580              2585              2590
Lys Gly Gly Lys Lys Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly
         2595              2600              2605
Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Ile Thr Gln Lys Leu
    2610              2615              2620
Pro Gln Ala Val Met Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala
2625              2630              2635              2640
Gln Arg Val Glu Tyr Leu Leu Lys Ala Trp Ala Glu Lys Lys Asp Pro
             2645              2650              2655
Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu
         2660              2665              2670
Arg Asp Ile Arg Thr Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro
    2675              2680              2685
Glu Glu Ala Arg Thr Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val
    2690              2695              2700
Gly Gly Pro Met Phe Asn Ser Lys Gly Gln Thr Cys Gly Tyr Arg Arg
2705              2710              2715              2720
Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr
             2725              2730              2735
Cys Tyr Val Lys Ala Leu Ala Ala Cys Lys Ala Ala Gly Ile Val Ala
         2740              2745              2750
Pro Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser
         2755              2760              2765
Gln Gly Thr Glu Glu Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala
    2770              2775              2780
Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr
2785              2790              2795              2800
Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu
             2805              2810              2815
Gly Pro Arg Gly Arg Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr
         2820              2825              2830
Pro Leu Ala Arg Ala Ala Trp Glu Thr Val Arg His Ser Pro Ile Asn
    2835              2840              2845
Ser Trp Leu Gly Asn Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg
    2850              2855              2860
Met Val Leu Met Thr His Phe Phe Ser Ile Leu Met Val Gln Asp Thr
2865              2870              2875              2880
```

26

```
Leu Asp Gln Asn Leu Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val
              2885                    2890                    2895
Asn Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp
          2900                    2905                    2910
Ala Phe Ser Met His Thr Tyr Ser His His Glu Leu Thr Arg Val Ala
        2915                    2920                    2925
Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser
      2930                    2935                    2940
Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala
2945                    2950                    2955                    2960
Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu
              2965                    2970                    2975
Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp
          2980                    2985                    2990
Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg
        2995                    3000                    3005
Ala Arg Pro Arg Ser Leu Leu Phe Gly Leu Leu Leu Leu Phe Val Gly
      3010                    3015                    3020
Val Gly Leu Phe Leu Leu Pro Ala Arg
3025                    3030
```

```
<210> 5
<211> 9674
<212> DNA
<213> Hepatitis C virus

<220>
<221> CDS
<222> (341)..(9442)

<400> 5
acccgcccct aataggggcg acactccgcc atgaatcact cccctgtgag gaactactgt 60

cttcacgcag aaagcgtcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120

ccccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg 180

aagactgggt cctttcttgg ataaacccac tctatgcccg gccatttggg cgtgcccccg 240

caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300

tgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat ccc  355
                                           Met Ser Thr Asn Pro
                                           1                   5
```

```
aaa cct caa aga aaa acc aaa aga aac act aac cgt cgc cca caa gac  403
Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp
              10                  15                  20

gtt aag ttt ccg ggc ggc ggc cag atc gtt ggc gga gta tac ttg ttg  451
Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu
          25                  30                  35

ccg cgc agg ggc ccc agg ttg ggt gtg cgc gcg aca agg aag gct tcg  499
Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr Arg Lys Ala Ser
          40                  45                  50
```

27

```
gag cgg tcc cag cca cgt ggg agg cgc cag ccc atc ccc aaa cat cgg    547
Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys His Arg
        55                  60                  65

cgc tcc act ggc aag tcc tgg ggg aag cca gga tac ccc tgg ccc ctg    595
Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly Tyr Pro Trp Pro Leu
 70                  75                  80                      85

tat ggg aat gag ggg ctc ggt tgg gca gga tgg ctc ctg tcc cct cga    643
Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg
                    90                  95                  100

ggt tcc cgt ccc tca tgg ggc ccc aat gac ccc cgg cat agg tcg cgc    691
Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro Arg His Arg Ser Arg
                105                 110                 115

aat gtg ggt aag gtc atc gat acc cta acg tgc ggc ttt gcc gac ctc    739
Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu
            120                 125                 130

ttg ggg tac gtc ccc gtc gta ggc gcc ccg ctt agt ggc gtt gcc agt    787
Leu Gly Tyr Val Pro Val Val Gly Ala Pro Leu Ser Gly Val Ala Ser
            135                 140                 145

gct ctc gcg cac ggc gtg aga gtc ctg gag gac ggg gtt aat ttt gca    835
Ala Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Phe Ala
150                 155                 160                 165

aca ggg aac tta cct ggt tgc tcc ttt tct atc ttc ttg ctg gcc cta    883
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu
                    170                 175                 180

ctg tcc tgc atc act act ccg gtc tct gct gtc caa gtg aag aac acc    931
Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Val Gln Val Lys Asn Thr
                185                 190                 195

agc aac gcc tat atg gcg act aac gac tgt tcc aat gac agc atc act    979
Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser Asn Asp Ser Ile Thr
            200                 205                 210

tgg cag ctt gag gcc gca gtc ctc cat gtc ccc ggg tgc gtc ccg tgc    1027
Trp Gln Leu Glu Ala Ala Val Leu His Val Pro Gly Cys Val Pro Cys
            215                 220                 225

gag aaa atg ggg aac aca tca cgg tgc tgg ata cca gtc tca cca aac    1075
Glu Lys Met Gly Asn Thr Ser Arg Cys Trp Ile Pro Val Ser Pro Asn
230                 235                 240                 245

gtg gct gtg cgg cag cct ggc gcc ctc acg cgg ggc ttg cgg acg cac    1123
Val Ala Val Arg Gln Pro Gly Ala Leu Thr Arg Gly Leu Arg Thr His
                250                 255                 260

atc gac atg gtc gtg ttg tcc gcc acg ctc tgc tcc gct ctc tac gtg    1171
Ile Asp Met Val Val Leu Ser Ala Thr Leu Cys Ser Ala Leu Tyr Val
                265                 270                 275
```

28

```
ggg gac ctc tgt ggc ggg gtg atg ctc gcg tcc cag atg ttc att gtc    1219
Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ser Gln Met Phe Ile Val
        280             285             290

tcg ccg cag cac cac tgg ttc gtg cag gaa tgc aat tgc tcc atc tac    1267
Ser Pro Gln His His Trp Phe Val Gln Glu Cys Asn Cys Ser Ile Tyr
        295             300             305

cct ggc gcc atc act ggg cac cgt atg gca tgg gac atg atg atg aac    1315
Pro Gly Ala Ile Thr Gly His Arg Met Ala Trp Asp Met Met Met Asn
310             315             320             325

tgg tcg ccc acg acc acc atg atc ctg gcg tac gtg atg cgc gtt ccc    1363
Trp Ser Pro Thr Thr Thr Met Ile Leu Ala Tyr Val Met Arg Val Pro
        330             335             340

gag gtc atc ata gac atc att agc gga gct cac tgg ggc gtc atg ttt    1411
Glu Val Ile Ile Asp Ile Ile Ser Gly Ala His Trp Gly Val Met Phe
        345             350             355

ggc ctg gcc tac ttc tct atg cag gga gcg tgg gcg aag gtc gtt gtc    1459
Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp Ala Lys Val Val Val
        360             365             370

atc ctc ctg ctg gcc tct ggg gtg gac gcg tac acc acc acg act ggg    1507
Ile Leu Leu Leu Ala Ser Gly Val Asp Ala Tyr Thr Thr Thr Thr Gly
        375             380             385

agc gct gct ggg cgc act acc agt agc ctg gcc agc gcc ttc tcc cct    1555
Ser Ala Ala Gly Arg Thr Thr Ser Ser Leu Ala Ser Ala Phe Ser Pro
390             395             400             405

ggc gct cgg cag aac att cag ctc att aat acc aat ggt agc tgg cac    1603
Gly Ala Arg Gln Asn Ile Gln Leu Ile Asn Thr Asn Gly Ser Trp His
        410             415             420

atc aac cgc acc gcc ctg aat tgc aac gat tcc ttg cac acc ggc ttc    1651
Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser Leu His Thr Gly Phe
        425             430             435

ttc acg gcc ctg ttc tac atc cat aag ttc aac tcg tcg gga tgt ccc    1699
Phe Thr Ala Leu Phe Tyr Ile His Lys Phe Asn Ser Ser Gly Cys Pro
        440             445             450

gag cgc ctg tcc gcc tgt cgc aac atc gag gac ttc cgg ata gga tgg    1747
Glu Arg Leu Ser Ala Cys Arg Asn Ile Glu Asp Phe Arg Ile Gly Trp
        455             460             465

ggc gcc ctg caa tac gac gac aat gtc acc aat cca gaa gat atg agg    1795
Gly Ala Leu Gln Tyr Asp Asp Asn Val Thr Asn Pro Glu Asp Met Arg
470             475             480             485

cca tat tgc tgg cac tac cca cca aaa cag tgt ggc gta gtc ccc gca    1843
Pro Tyr Cys Trp His Tyr Pro Pro Lys Gln Cys Gly Val Val Pro Ala
            490             495             500

ggg acc gtg tgc ggc cca gtg tac tgt ttc acc cct agc ccg gtg gta    1891
```

29

```
Gly Thr Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val
            505             510             515

gtg ggc acg acc gat aga ctt gga gtg cct act tac acg tgg gga gag    1939
Val Gly Thr Thr Asp Arg Leu Gly Val Pro Thr Tyr Thr Trp Gly Glu
            520             525             530

aat gag aca gat gtc ttc cta ttg aac agc acc cga cca ccg tcg ggg    1987
Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr Arg Pro Pro Ser Gly
            535             540             545

tca tgg ttt ggc tgc acg tgg atg aac tcc act ggc ttc acc aag acc    2035
Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe Thr Lys Thr
550             555             560             565

tgc ggc gca cca ccc tgc cgc act aga gct gac ttc aat acc agc aca    2083
Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp Phe Asn Thr Ser Thr
            570             575             580

gat ctg ttg tgc ccc acg gac tgt ttt aga aaa cat cct gaa gcc act    2131
Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr
            585             590             595

tac atc aaa tgt ggt tcc ggg cct tgg ctc acg cca aag tgt ctg gtt    2179
Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr Pro Lys Cys Leu Val
            600             605             610

gac tac ccc tac agg ctc tgg cat tac cct tgc aca gtc aat tac tcc    2227
Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Val Asn Tyr Ser
            615             620             625

acc ttc aag atc agg atg tat gtg ggg gga gtt gag cac agg ctc atg    2275
Thr Phe Lys Ile Arg Met Tyr Val Gly Gly Val Glu His Arg Leu Met
630             635             640             645

gcc gcg tgc aat ttc act cgt ggg gat cgc tgc aac ttg gag gat agg    2323
Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys Asn Leu Glu Asp Arg
            650             655             660

gac aga agt caa cag act cct ctg ttg cac tcc acc acg gaa tgg gcc    2371
Asp Arg Ser Gln Gln Thr Pro Leu Leu His Ser Thr Thr Glu Trp Ala
            665             670             675

att ttg ccc tgc tct ttc tca gac ttg ccc gct ttg tcg act ggt ctt    2419
Ile Leu Pro Cys Ser Phe Ser Asp Leu Pro Ala Leu Ser Thr Gly Leu
            680             685             690

ctc cac ctc cac caa aat atc gtg gac gta caa tat atg tat ggc ctg    2467
Leu His Leu His Gln Asn Ile Val Asp Val Gln Tyr Met Tyr Gly Leu
            695             700             705

tca cct gcc ctc aca caa tat atc gtt cga tgg gag tgg gta gta ctc    2515
Ser Pro Ala Leu Thr Gln Tyr Ile Val Arg Trp Glu Trp Val Val Leu
710             715             720             725

tta ttc ctg ctc cta gcg gac gcc agg gtc tgc gcc tgc ttg tgg atg    2563
Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ala Cys Leu Trp Met
```

30

ctc atc ttg ctg ggc caa gcc gaa gca gca ctg gag aag ctg gtc gtc   2611
Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu Glu Lys Leu Val Val
            745   750   755

ttg cac gct gcg agc gca gct agc tgc aat ggc ttc ctg tat ttt gtc   2659
Leu His Ala Ala Ser Ala Ala Ser Cys Asn Gly Phe Leu Tyr Phe Val
            760   765   770

atc ttt ctc gtg gct gct tgg cac atc aag ggt agg gtg gtc ccc ttg   2707
Ile Phe Leu Val Ala Ala Trp His Ile Lys Gly Arg Val Val Pro Leu
            775   780   785

gct gct tat tcc ctt act ggc ctg tgg ccg ttc tgc cta ctg ctc cta   2755
Ala Ala Tyr Ser Leu Thr Gly Leu Trp Pro Phe Cys Leu Leu Leu Leu
790   795   800   805

gca ctg ccc cag cag gct tac gcc tat gat gca tct gtg cac gga cag   2803
Ala Leu Pro Gln Gln Ala Tyr Ala Tyr Asp Ala Ser Val His Gly Gln
                    810   815   820

gtg ggc gcg gct ttg cta gta ctg att acc ctc ttt aca ctc acc ccg   2851
Val Gly Ala Ala Leu Leu Val Leu Ile Thr Leu Phe Thr Leu Thr Pro
            825   830   835

ggg tat aag acc ctt ctc agc cag tcc ctg tgg tgg ttg tgc tat ctc   2899
Gly Tyr Lys Thr Leu Leu Ser Gln Ser Leu Trp Trp Leu Cys Tyr Leu
            840   845   850

ctg acc ctg gcg gaa acc atg gtc cag gag tgg gca cca tcc atg cag   2947
Leu Thr Leu Ala Glu Thr Met Val Gln Glu Trp Ala Pro Ser Met Gln
            855   860   865

gcg cgc ggc ggc cgt gat ggc atc ata tgg gcc gcc acc ata ttt tgc   2995
Ala Arg Gly Gly Arg Asp Gly Ile Ile Trp Ala Ala Thr Ile Phe Cys
870   875   880   885

ccg ggc gta gtg ttt gac ata acc aag tgg ctc tta gcg gtg ctt ggg   3043
Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu Leu Ala Val Leu Gly
                    890   895   900

cct ggt tac ctc cta aga ggt gct ttg acg cgc gtg cca tat ttc gtc   3091
Pro Gly Tyr Leu Leu Arg Gly Ala Leu Thr Arg Val Pro Tyr Phe Val
                    905   910   915

aga gcc cac gct ctg ctg aga atg tgc act atg gtg agg cac ctc gcg   3139
Arg Ala His Ala Leu Leu Arg Met Cys Thr Met Val Arg His Leu Ala
            920   925   930

ggg ggt agg tac gtc cag atg gcg cta tta gcc ctt ggc agg tgg act   3187
Gly Gly Arg Tyr Val Gln Met Ala Leu Leu Ala Leu Gly Arg Trp Thr
            935   940   945

ggc act tac atc tat gac cac ctc acc cct atg ·tcg gat tgg gct gct   3235
Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met Ser Asp Trp Ala Ala
950   955   960   965

31

```
agc ggc ctg cgg gac ttg gcg gtc gct gtg gag cct atc atc ttc agt      3283
Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Ile Ile Phe Ser
                970                 975                 980

ccg atg gag aag aaa gtc atc gtt tgg gga gcg gag acg gct gcg tgc      3331
Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala Glu Thr Ala Ala Cys
            985                 990                 995

ggg gac atc ttg cac gga ctt ccc gtg tcc gcc cga ctc ggt cgg gag      3379
Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala Arg Leu Gly Arg Glu
        1000                1005                1010

atc ctc ctt ggc cca gct gat ggc tac acc tcc aag ggg tgg aag ctt      3427
Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser Lys Gly Trp Lys Leu
    1015                1020                1025

ctc gcc ccc atc acc gct tac gcc cag cag aca cga ggt ctc ttg ggc      3475
Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu Gly
1030                1035                1040                1045

tct ata gtg gtg agc atg acg ggg cgt gac aag aca gaa cag gcc ggg      3523
Ser Ile Val Val Ser Met Thr Gly Arg Asp Lys Thr Glu Gln Ala Gly
                1050                1055                1060

gag gtc caa gtc ctg tcc aca gtc act cag tcc ttc ctc gga aca tcc      3571
Glu Val Gln Val Leu Ser Thr Val Thr Gln Ser Phe Leu Gly Thr Ser
                1065                1070                1075

att tcg ggg gtc tta tgg act gtt tac cac gga gct ggc aac aag aca      3619
Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly Ala Gly Asn Lys Thr
            1080                1085                1090

cta gcc ggc tcg cgg ggc ccg gtc acg cag atg tac tcg agc gcc gag      3667
Leu Ala Gly Ser Arg Gly Pro Val Thr Gln Met Tyr Ser Ser Ala Glu
        1095                1100                1105

ggg gac ttg gtc ggg tgg ccc agc cct cct ggg acc aaa tct ttg gag      3715
Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly Thr Lys Ser Leu Glu
1110                1115                1120                1125

ccg tgt acg tgt gga gcg gtc gac ctg tat ttg gtc acg cgg aac gct      3763
Pro Cys Thr Cys Gly Ala Val Asp Leu Tyr Leu Val Thr Arg Asn Ala
                1130                1135                1140

gat gtc atc ccg gct cga aga cgc ggg gac aag cgg gga gcg ctg ctc      3811
Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys Arg Gly Ala Leu Leu
            1145                1150                1155

tcc ccg aga ccc ctt tcg acc ttg aag ggg tcc tcg ggg gga cct gtg      3859
Ser Pro Arg Pro Leu Ser Thr Leu Lys Gly Ser Ser Gly Gly Pro Val
            1160                1165                1170

ctt tgc cct agg ggc cac gct gtc gga atc ttc cgg gca gct gtg tgc      3907
Leu Cys Pro Arg Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys
    1175                1180                1185
```

```
tct cgg ggt gtg gct aag tcc ata gat ttc atc ccc gtt gag acg ctc     3955
Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile Pro Val Glu Thr Leu
1190            1195            1200            1205

gac atc gtc acg cgg tct ccc acc ttt agt gac aac agc aca cca cca     4003
Asp Ile Val Thr Arg Ser Pro Thr Phe Ser Asp Asn Ser Thr Pro Pro
            1210            1215            1220

gct gtg ccc cag acc tat cag gtg ggg tac ttg cac gcc ccc act ggc     4051
Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu His Ala Pro Thr Gly
        1225            1230            1235


agt gga aaa agc acc aag gtc ccc gtc gcg tac gcc gcc cag ggg tat     4099
Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr Ala Ala Gln Gly Tyr
     1240            1245            1250

aaa gtg ctg gtg ctc aat ccc tcg gtg gct gcc acc ctg gga ttt ggg     4147
Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly
     1255            1260            1265

gcg tac ttg tcc aag gca cat ggc atc aac ccc aac att agg act gga     4195
Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro Asn Ile Arg Thr Gly
1270            1275            1280            1285

gtc aga act gtg acg acc ggg gag ccc att aca tac tcc acg tat ggt     4243
Val Arg Thr Val Thr Thr Gly Glu Pro Ile Thr Tyr Ser Thr Tyr Gly
            1290            1295            1300

aaa ttc ctc gcc gat ggg ggc tgc gca ggc ggc gcc tat gac atc atc     4291
Lys Phe Leu Ala Asp Gly Gly Cys Ala Gly Gly Ala Tyr Asp Ile Ile
        1305            1310            1315

ata tgc gat gaa tgc cac tct gtg gat gct acc act att ctc ggc atc     4339
Ile Cys Asp Glu Cys His Ser Val Asp Ala Thr Thr Ile Leu Gly Ile
        1320            1325            1330

ggg aca gtc ctt gac caa gca gag aca gcc ggg gtc agg cta act gta     4387
Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Val Arg Leu Thr Val
     1335            1340            1345

ctg gcc acg gcc acg ccc ccc ggg tcg gtg aca acc ccc cat ccc aat     4435
Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Thr Pro His Pro Asn
1350            1355            1360            1365

ata gag gag gta gcc ctc gga cag gag ggt gag atc ccc ttc tat ggg     4483
Ile Glu Glu Val Ala Leu Gly Gln Glu Gly Glu Ile Pro Phe Tyr Gly
            1370            1375            1380

agg gcg ttt ccc ctg tct tac atc aag gga ggg agg cac ttg att ttc     4531
Arg Ala Phe Pro Leu Ser Tyr Ile Lys Gly Gly Arg His Leu Ile Phe
        1385            1390            1395

tgc cac tca aag aaa aag tgt gac gag ctc gca acg gcc ctt cgg ggc     4579
Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Thr Ala Leu Arg Gly
        1400            1405            1410
```

33

```
atg ggc ttg aac gct gtg gca tat tac aga ggg ttg gac gtc tcc ata       4627
Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Ile
    1415                1420            1425


ata cca act caa gga gat gtg gtg gtc gtt gcc acc gac gcc ctc atg       4675
Ile Pro Thr Gln Gly Asp Val Val Val Ala Thr Asp Ala Leu Met
1430                1435            1440            1445


acg ggg tat act gga gac ttt gac tcc gtg atc gac tgc aac gta gcg       4723
Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Val Ala
            1450            1455            1460


gtc acc cag gcc gta gac ttc agc ctg gac ccc acc ttc act ata acc       4771
Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile Thr
        1465            1470            1475


aca cag act gtc ccg caa gac gct gtc tca cgt agt cag cgc cga ggg       4819
Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg Ser Gln Arg Arg Gly
        1480            1485            1490


cgc acg ggt aga gga aga ctg ggc att tat agg tat gtt tcc act ggt       4867
Arg Thr Gly Arg Gly Arg Leu Gly Ile Tyr Arg Tyr Val Ser Thr Gly
    1495            1500            1505


gag cga gcc tca gga atg ttt gac agt gta gta ctc tgt gag tgc tac       4915
Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val Leu Cys Glu Cys Tyr
1510            1515            1520            1525


gac gca gga gct gct tgg tat gag ctc tca cca gtg gag acg acc gtc       4963
Asp Ala Gly Ala Ala Trp Tyr Glu Leu Ser Pro Val Glu Thr Thr Val
            1530            1535            1540


agg ctc agg gcg tat ttc aac acg cct ggc ttg cct gtg tgc cag gac       5011
Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu Pro Val Cys Gln Asp
        1545            1550            1555


cac ctt gag ttt tgg gag gca gtt ttc acc ggc ctc aca cac ata gac       5059
His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly Leu Thr His Ile Asp
        1560            1565            1570


gct cat ttc ctt tcc cag aca aag cag tcg ggg gaa aat ttc gca tac       5107
Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Phe Ala Tyr
    1575            1580            1585


tta gta gcc tat cag gcc aca gtg tgc gcc agg gcc aaa gcg ccc ccc       5155

Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Lys Ala Pro Pro
1590            1595            1600            1605


ccg tcc tgg gac gtc atg tgg aag tgc ttg act cga ctc aag ccc acg       5203
Pro Ser Trp Asp Val Met Trp Lys Cys Leu Thr Arg Leu Lys Pro Thr
            1610            1615            1620


ctt gtg ggc cct aca cct ctc ctg tac cgt ttg ggc tct gtt acc aac       5251
Leu Val Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ser Val Thr Asn
        1625            1630            1635
```

34

```
gag gtc acc ctt aca cac ccc gtg aca aaa tac atc gcc aca tgc atg        5299
Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr Ile Ala Thr Cys Met
        1640                1645                1650

caa gct gac ctc gag gtc atg acc agc acg tgg gtc ctg gct ggg gga        5347
Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp Val Leu Ala Gly Gly
    1655                1660                1665

gtc tta gca gcc gtc gcc gcg tat tgc tta gcg acc ggg tgt gtt tcc        5395
Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala Thr Gly Cys Val Ser
1670                1675                1680                1685

atc att ggc cgt tta cac atc aac cag cga gct gtc gtc gct ccg gac        5443
Ile Ile Gly Arg Leu His Ile Asn Gln Arg Ala Val Val Ala Pro Asp
            1690                1695                1700

aag gag gtc ctc tat gag gct ttt gat gag atg gag gaa tgt gcc tcc        5491
Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met Glu Glu Cys Ala Ser
            1705                1710                1715

aga gcg gct ctc ctt gaa gag ggg cag cgg ata gcc gag atg ctg aag        5539
Arg Ala Ala Leu Leu Glu Glu Gly Gln Arg Ile Ala Glu Met Leu Lys
        1720                1725                1730

tcc aag atc caa ggc tta ttg cag caa gcc tct aaa cag gcc cag gac        5587
Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser Lys Gln Ala Gln Asp
    1735                1740                1745

ata caa ccc gct gtg caa gct tcg tgg ccc aag atg gag caa ttc tgg        5635
Ile Gln Pro Ala Val Gln Ala Ser Trp Pro Lys Met Glu Gln Phe Trp
1750                1755                1760                1765

gcc aaa cat atg tgg aac ttc ata agc ggc att cag tac ctc gca gga        5683
Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala Gly
            1770                1775                1780

ctg tca aca ctg cca ggg aac cct gct gtg gct tcc atg atg gca ttc        5731
Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala Ser Met Met Ala Phe
            1785                1790                1795

agc gcc gcc ctc acc agt ccg ttg tca act agc acc acc atc ctt ctt        5779
Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser Thr Thr Ile Leu Leu
        1800                1805                1810

aac att ctg ggg ggc tgg ctg gcg tcc caa att gcg cca ccc gcg ggg        5827
Asn Ile Leu Gly Gly Trp Leu Ala Ser Gln Ile Ala Pro Pro Ala Gly
    1815                1820                1825

gcc act ggc ttt gtt gtc agt ggc ctg gtg gga gct gct gtt ggc agc        5875
Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly Ala Ala Val Gly Ser
1830                1835                1840                1845

ata ggc ttg ggt aaa gtg ctg gtg gac atc ctg gca ggg tat ggt gcg        5923
Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly Ala
            1850                1855                1860

ggc att tcg ggg gcc ctc gtc gcg ttt aag atc atg tct ggc gag aag        5971
```

35

```
Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Lys
         1865                1870                1875


ccc tcc atg gag gat gtc atc aac ttg ctg cct ggg att ctg tct cca    6019
Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro Gly Ile Leu Ser Pro
         1880                1885                1890


ggt gct ctg gtg gtg gga gtc atc tgc gcg gcc att ctg cgc cgc cat    6067
Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala Ile Leu Arg Arg His
    1895                1900                1905


gtg gga ccg ggg gaa ggc gcg gtc caa tgg atg aac agg ctt atc gcc    6115
Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala
1910                1915                1920                1925


ttc gct tcc aga gga aac cac gtc gcc cct act cac tac gtg acg gag    6163
Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr His Tyr Val Thr Glu
             1930                1935                1940


tcg gat gcg tcg cag cgt gtc acc caa ctg ctt ggc tct ctc act ata    6211
Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu Gly Ser Leu Thr Ile
         1945                1950                1955


act agt cta ctc agg aga ctt cac aac tgg atc act gag gat tgc ccc    6259
Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile Thr Glu Asp Cys Pro
         1960                1965                1970


atc cca tgc gcc ggc tcg tgg ctc cgc gat gtg tgg gac tgg gtc tgt    6307
Ile Pro Cys Ala Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Val Cys
    1975                1980                1985


acc atc cta aca gac ttt aag aac tgg ctg acc tcc aag ctg ttc cca    6355
Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr Ser Lys Leu Phe Pro
1990                1995                2000                2005


aag atg cct ggc ctc ccc ttt atc tct tgc caa aag ggg tac aag ggc    6403
Lys Met Pro Gly Leu Pro Phe Ile Ser Cys Gln Lys Gly Tyr Lys Gly
             2010                2015                2020


gtg tgg gcc ggc act ggc atc atg acc aca cga tgc ccc tgc ggc gcc    6451
Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg Cys Pro Cys Gly Ala
         2025                2030                2035


aac atc tct ggc aac gtc cgc ttg ggc tct atg aga atc aca gga ccc    6499
Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met Arg Ile Thr Gly Pro
         2040                2045                2050


aaa acc tgc atg aac acc tgg cag ggg acc ttt cct atc aat tgt tat    6547
Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe Pro Ile Asn Cys Tyr
    2055                2060                2065


aca gaa ggc cag tgc ttg ccg aaa ccc gcg tta aac ttc aag acc gcc    6595
Thr Glu Gly Gln Cys Leu Pro Lys Pro Ala Leu Asn Phe Lys Thr Ala
2070                2075                2080                2085


atc tgg aga gtg gcg gcc tca gag tac gcg gaa gtg acg cag cac gga    6643
```

36

```
Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu Val Thr Gln His Gly
            2090                    2095                    2100

tca tat gcc tat ata aca ggg ctg acc act gac aac tta aaa gtc cct    6691
Ser Tyr Ala Tyr Ile Thr Gly Leu Thr Thr Asp Asn Leu Lys Val Pro
            2105                    2110                    2115

tgc caa ctc ccc tct cca gag ttt ttc tct tgg gtg gac gga gta caa    6739
Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp Val Asp Gly Val Gln
            2120                    2125                    2130

atc cat agg tcc gcc ccc aca cca aag ccg ttt ttc cgg gat gag gtc    6787
Ile His Arg Ser Ala Pro Thr Pro Lys Pro Phe Phe Arg Asp Glu Val
           2135                    2140                    2145

tcg ttc agc gtt ggg ctc aat tca ttt gtc gtc ggg tct cag ctt ccc    6835
Ser Phe Ser Val Gly Leu Asn Ser Phe Val Val Gly Ser Gln Leu Pro
2150                    2155                    2160                    2165

tgt gac cct gag ccc gac act gag gta gtg atg tcc atg cta aca gac    6883
Cys Asp Pro Glu Pro Asp Thr Glu Val Val Met Ser Met Leu Thr Asp
                  2170                    2175                    2180

cca tcc cat atc acg gcg gag gct gca gcg cgg cgt tta gcg cgg ggg    6931
Pro Ser His Ile Thr Ala Glu Ala Ala Ala Arg Arg Leu Ala Arg Gly
                  2185                    2190                    2195

tca ccc cca tct gag gca agc tcc tca gcg agc cag ctg tcg gcg cca    6979
Ser Pro Pro Ser Glu Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala Pro
            2200                    2205                    2210

tcg ctg cga gcc acc tgc acc acc cac ggt agg acc tat gat gtg gac    7027
Ser Leu Arg Ala Thr Cys Thr Thr His Gly Arg Thr Tyr Asp Val Asp
      2215                    2220                    2225

atg gtg gat gcc aac ctg ttc atg ggg ggc ggc gtg att cgg ata gag    7075
Met Val Asp Ala Asn Leu Phe Met Gly Gly Gly Val Ile Arg Ile Glu
2230                    2235                    2240                    2245

tct gag tcc aaa gtg gtc gtt ctg gac tcc ctc gac tca atg acc gag    7123
Ser Glu Ser Lys Val Val Val Leu Asp Ser Leu Asp Ser Met Thr Glu
            2250                    2255                    2260

gaa gag ggc gac ctt gag cct tca gta cca tcg gag tat atg ctc ccc    7171
Glu Glu Gly Asp Leu Glu Pro Ser Val Pro Ser Glu Tyr Met Leu Pro
            2265                    2270                    2275

agg aag agg ttc cca ccg gcc tta ccg gct tgg gcg cgg cct gat tac    7219
Arg Lys Arg Phe Pro Pro Ala Leu Pro Ala Trp Ala Arg Pro Asp Tyr
            2280                    2285                    2290

aac cca ccg ctt gtg gaa tcg tgg aag agg cca gat tac caa cca ccc    7267
Asn Pro Pro Leu Val Glu Ser Trp Lys Arg Pro Asp Tyr Gln Pro Pro
      2295                    2300                    2305

act gtt gcg ggc tgt gct ctc ccc ccc ccc aaa aag acc ccg acg cct    7315
Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys Lys Thr Pro Thr Pro
```

37

```
        2310                    2315                    2320                    2325

cct cca agg aga cgc cgg aca gtg ggt ctg agc gag agc acc ata gga    7363
Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser Glu Ser Thr Ile Gly
                2330                    2335                    2340


gat gcc ctc caa cag ctg gcc atc aag tcc ttt ggc cag ccc ccc cca    7411
Asp Ala Leu Gln Gln Leu Ala Ile Lys Ser Phe Gly Gln Pro Pro Pro
            2345                    2350                    2355


agc ggc gat tca ggc ctt tcc acg ggg gcg gac gcc gcc gac tcc ggc    7459
Ser Gly Asp Ser Gly Leu Ser Thr Gly Ala Asp Ala Ala Asp Ser Gly
            2360                    2365                    2370


gat cgg aca ccc cct gac gag ttg gct ctt tcg gag aca ggt tct acc    7507
Asp Arg Thr Pro Pro Asp Glu Leu Ala Leu Ser Glu Thr Gly Ser Thr
        2375                    2380                    2385


tcc tcc atg ccc ccc ctc gag ggg gag cct ggg gac cca gac ctg gag    7555
Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Glu
    2390                    2395                    2400                    2405


cct gag cag gta gag ctt caa cct cct ccc cag ggg ggg gag gca gct    7603
Pro Glu Gln Val Glu Leu Gln Pro Pro Pro Gln Gly Gly Glu Ala Ala
                2410                    2415                    2420


ccc ggc tcg gac tcg ggg tcc tgg tct act tgc tcc gag gag gat gac    7651
Pro Gly Ser Asp Ser Gly Ser Trp Ser Thr Cys Ser Glu Glu Asp Asp
            2425                    2430                    2435


tcc gtc gtg tgc tgc tcc atg tca tat tcc tgg acc ggg gct cta ata    7699
Ser Val Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Ile
            2440                    2445                    2450


act cct tgt agc ccc gaa gag gaa aag ttg cca att aac tcc ttg agc    7747
Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro Ile Asn Ser Leu Ser
        2455                    2460                    2465


aac tcg ctg ttg cga tac cat aac aag gta tac tgt act aca tca aag    7795
Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr Cys Thr Thr Ser Lys
    2470                    2475                    2480                    2485


agt gcc tca cta agg gct aaa aag gta act ttt gat agg atg caa gtg    7843
Ser Ala Ser Leu Arg Ala Lys Lys Val Thr Phe Asp Arg Met Gln Val
                2490                    2495                    2500


ctc gac gcc tat tat gat tca gtc tta aag gac atc aag cta gcg gcc    7891
Leu Asp Ala Tyr Tyr Asp Ser Val Leu Lys Asp Ile Lys Leu Ala Ala
                2505                    2510                    2515


tcc aag gtc agc gca agg ctc ctc acc tta gag gag gcg tgc caa ttg    7939
Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu Glu Ala Cys Gln Leu
            2520                    2525                    2530


acc cca ccc cac tct gca aga tcc aag tat ggg ttt ggg gct aag gag    7987
Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly Phe Gly Ala Lys Glu
        2535                    2540                    2545
```

38

```
gtc cgc agc ttg tcc ggg agg gcc gtc aac cac atc aag tcc gtg tgg     8035
Val Arg Ser Leu Ser Gly Arg Ala Val Asn His Ile Lys Ser Val Trp
2550            2555             2560             2565

aag gac ctc ttg gaa gac tca caa aca cca att cct aca acc atc atg     8083
Lys Asp Leu Leu Glu Asp Ser Gln Thr Pro Ile Pro Thr Thr Ile Met
            2570             2575             2580

gcc aaa aat gag gtg ttc tgc gtg gac ccc gcc aag ggg ggt aaa aaa     8131
Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala Lys Gly Gly Lys Lys
        2585             2590             2595

cca gct cgc ctt atc gtt tac cct gac ctc ggc gtc agg gtc tgc gag     8179
Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu
        2600             2605             2610

aag atg gcc ctt tat gat gtc aca caa aag ctt cct cag gcg gtg atg     8227
Lys Met Ala Leu Tyr Asp Val Thr Gln Lys Leu Pro Gln Ala Val Met
    2615             2620             2625

ggg gct tct tat ggc ttc cag tac tcc ccc gct cag cgg gtg gag ttt     8275
Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Glu Phe
2630             2635             2640             2645

ctc ttg aag gca tgg gcg gaa aag aga gac cct atg ggt ttt tcg tat     8323
Leu Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro Met Gly Phe Ser Tyr
            2650             2655             2660

gat acc cga tgc ttt gac tca acc gtc act gag aga gac atc agg act     8371
Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Arg Asp Ile Arg Thr
            2665             2670             2675

gag gag tcc ata tac cag gcc tgc tcc tta ccc gag gag gcc cga act     8419
Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro Glu Glu Ala Arg Thr
        2680             2685             2690

gcc ata cac tcg ctg act gag aga ctc tat gtg gga ggg ccc atg ttc     8467
Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Met Phe
    2695             2700             2705

aac agc aag ggc cag tcc tgc ggg tac agg cgt tgc cgc gcc agc ggg     8515
Asn Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly
2710             2715             2720             2725

gtg ctt acc act agt atg ggg aac acc atc aca tgc tat gta aaa gcc     8563
Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr Cys Tyr Val Lys Ala
            2730             2735             2740

cta gcg gct tgc aag gct gcg ggg ata att gcg ccc acg atg ctg gta     8611
Leu Ala Ala Cys Lys Ala Ala Gly Ile Ile Ala Pro Thr Met Leu Val
            2745             2750             2755

tgc ggc gac gac ttg gtc gtc atc tca gaa agc cag ggg act gag gag     8659
Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser Gln Gly Thr Glu Glu
        2760             2765             2770
```

39

```
gac gag cgg aac ctg aga gcc ttc acg gag gct atg acc agg tat tct    8707
Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser
   2775               2780               2785

gcc cct cct ggt gac ccc ccc aga ccg gaa tat gac ctg gag cta ata    8755
Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr Asp Leu Glu Leu Ile
2790               2795               2800               2805

aca tct tgt tcc tca aac gtg tct gtg gca ctt ggc cca cag ggc cgc    8803
Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu Gly Pro Gln Gly Arg
            2810               2815               2820

cgc aga tac tac ctg acc aga gac ccc acc act tca att gcc cgg gct    8851
Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Ser Ile Ala Arg Ala
            2825               2830               2835

gcc tgg gaa aca gtt aga cac tcc cct gtc aat tca tgg ctg gga aac    8899
Ala Trp Glu Thr Val Arg His Ser Pro Val Asn Ser Trp Leu Gly Asn
         2840               2845               2850

atc atc cag tac gct cca acc ata tgg gtt cgc atg gtc ctg atg aca    8947
Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg Met Val Leu Met Thr
   2855               2860               2865

cac ttc ttc tcc att ctc atg gcc cag gac acc cta gac cag aac ctt    8995
His Phe Phe Ser Ile Leu Met Ala Gln Asp Thr Leu Asp Gln Asn Leu
2870               2875               2880               2885

aac ttt gaa atg tac gga tcg gtg tac tcc gtg agt cct ctg gac ctc    9043
Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Ser Pro Leu Asp Leu
            2890               2895               2900

cca gcc ata att gaa agg tta cac ggg ctt gac gcc ttc tct ctg cac    9091
Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp Ala Phe Ser Leu His
            2905               2910               2915

aca tac act ccc cac gaa ctg acg cgg gtg gct tca gcc ctc aga aaa    9139
Thr Tyr Thr Pro His Glu Leu Thr Arg Val Ala Ser Ala Leu Arg Lys
         2920               2925               2930

ctt ggg gcg cca ccc ctc aga gcg tgg aag agt cgg gcg cgt gca gtt    9187
Leu Gly Ala Pro Pro Leu Arg Ala Trp Lys Ser Arg Ala Arg Ala Val
   2935               2940               2945

agg gcg tcc ctc atc tcc cgt ggg ggg agg gcg gcc gtt tgc ggt cgg    9235
Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala Ala Val Cys Gly Arg
2950               2955               2960               2965

tac ctc ttc aac tgg gcg gtg aag acc aag ctc aaa ctc act cct ttg    9283
Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu
            2970               2975               2980

ccg gag gca cgc ctc ctg gat ttg tcc agt tgg ttt acc gtc ggc gcc    9331
Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Gly Ala
            2985               2990               2995

ggc ggg ggc gac att tat cac agc gtg tcg cgt gcc cga ccc cgc cta    9379
```

40

```
Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg Ala Arg Pro Arg Leu
     3000                 3005                 3010

tta ctc ctt agc cta ctc cta ctt tct gta ggg gta ggc ctc ttc cta    9427
Leu Leu Leu Ser Leu Leu Leu Leu Ser Val Gly Val Gly Leu Phe Leu
   3015                 3020                 3025

ctc ccc gct cga tag agcggcacac attagctaca ctccatagct aactgttcct    9482
Leu Pro Ala Arg
3030

tttttttttt tttttttttt tttttttttt ttttttttctt tttttttttt tttccctctt 9542

tcttcccttc tcatcttatt ctactttctt tcttggtggc tccatcttag ccctggtcac 9602

ggctagctgt gaaaggtccg tgagccgcat gactgcagag agtgccgtaa ctggtctctc 9662

tgcagatcat gt                                                      9674


<210> 6
<211> 3033
<212> PRT
<213> Hepatitis C virus

<400> 6
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
 1               5                  10                   15
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gln Ile Val Gly
           20             25                 30
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
         35                 40                 45
Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
     50                 55                 60
Ile Pro Lys His Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly
 65                 70                 75                 80
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
                 85                 90                 95
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro
             100                105                110
Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
         115                120                125
Gly Phe Ala Asp Leu Leu Gly Tyr Val Pro Val Val Gly Ala Pro Leu
     130                135                140
Ser Gly Val Ala Ser Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
145                150                155                160
Gly Val Asn Phe Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
                 165                170                175
Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Val
             180                185                190
Gln Val Lys Asn Thr Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser
         195                200                205
Asn Asp Ser Ile Thr Trp Gln Leu Glu Ala Ala Val Leu His Val Pro
     210                215                220
Gly Cys Val Pro Cys Glu Lys Met Gly Asn Thr Ser Arg Cys Trp Ile
225                230                235                240
Pro Val Ser Pro Asn Val Ala Val Arg Gln Pro Gly Ala Leu Thr Arg
```

```
                    245                     250                     255
Gly Leu Arg Thr His Ile Asp Met Val Val Leu Ser Ala Thr Leu Cys
        260                     265                     270
Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ser
        275                     280                     285
Gln Met Phe Ile Val Ser Pro Gln His Trp Phe Val Gln Glu Cys
    290                     295                     300
Asn Cys Ser Ile Tyr Pro Gly Ala Ile Thr Gly His Arg Met Ala Trp
305                     310                     315                     320
Asp Met Met Met Asn Trp Ser Pro Thr Thr Thr Met Ile Leu Ala Tyr
                325                     330                     335
Val Met Arg Val Pro Glu Val Ile Ile Asp Ile Ile Ser Gly Ala His
        340                     345                     350
Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp
        355                     360                     365
Ala Lys Val Val Val Ile Leu Leu Leu Ala Ser Gly Val Asp Ala Tyr
    370                     375                     380
Thr Thr Thr Thr Gly Ser Ala Ala Gly Arg Thr Thr Ser Ser Leu Ala
385                     390                     395                     400
Ser Ala Phe Ser Pro Gly Ala Arg Gln Asn Ile Gln Leu Ile Asn Thr
                405                     410                     415
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
                420                     425                     430
Leu His Thr Gly Phe Phe Thr Ala Leu Phe Tyr Ile His Lys Phe Asn
        435                     440                     445
Ser Ser Gly Cys Pro Glu Arg Leu Ser Ala Cys Arg Asn Ile Glu Asp
    450                     455                     460
Phe Arg Ile Gly Trp Gly Ala Leu Gln Tyr Asp Asp Asn Val Thr Asn
465                     470                     475                     480
Pro Glu Asp Met Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Gln Cys
                485                     490                     495
Gly Val Val Pro Ala Gly Thr Val Cys Gly Pro Val Tyr Cys Phe Thr
        500                     505                     510
Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Leu Gly Val Pro Thr
        515                     520                     525
Tyr Thr Trp Gly Glu Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr
    530                     535                     540
Arg Pro Pro Ser Gly Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr
545                     550                     555                     560
Gly Phe Thr Lys Thr Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp
                565                     570                     575
Phe Asn Thr Ser Thr Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys
        580                     585                     590
His Pro Glu Ala Thr Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr
    595                     600                     605
Pro Lys Cys Leu Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys
        610                     615                     620
Thr Val Asn Tyr Ser Thr Phe Lys Ile Arg Met Tyr Val Gly Gly Val
625                     630                     635                     640
Glu His Arg Leu Met Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys
                645                     650                     655
Asn Leu Glu Asp Arg Asp Arg Ser Gln Gln Thr Pro Leu Leu His Ser
                660                     665                     670
Thr Thr Glu Trp Ala Ile Leu Pro Cys Ser Phe Ser Asp Leu Pro Ala
        675                     680                     685
Leu Ser Thr Gly Leu Leu His Leu His Gln Asn Ile Val Asp Val Gln
        690                     695                     700
```

42

```
Tyr Met Tyr Gly Leu Ser Pro Ala Leu Thr Gln Tyr Ile Val Arg Trp
705                     710             715                     720
Glu Trp Val Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys
                725             730             735
Ala Cys Leu Trp Met Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu
            740             745             750
Glu Lys Leu Val Val Leu His Ala Ala Ser Ala Ala Ser Cys Asn Gly
        755             760             765
Phe Leu Tyr Phe Val Ile Phe Leu Val Ala Ala Trp His Ile Lys Gly
        770             775             780
Arg Val Val Pro Leu Ala Ala Tyr Ser Leu Thr Gly Leu Trp Pro Phe
785             790             795                     800
Cys Leu Leu Leu Ala Leu Pro Gln Gln Ala Tyr Ala Tyr Asp Ala
            805             810             815
Ser Val His Gly Gln Val Gly Ala Ala Leu Leu Val Leu Ile Thr Leu
            820             825             830
Phe Thr Leu Thr Pro Gly Tyr Lys Thr Leu Leu Ser Gln Ser Leu Trp
        835             840             845
Trp Leu Cys Tyr Leu Leu Thr Leu Ala Glu Thr Met Val Gln Glu Trp
        850             855             860
Ala Pro Ser Met Gln Ala Arg Gly Gly Arg Asp Gly Ile Ile Trp Ala
865             870             875                     880
Ala Thr Ile Phe Cys Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu
            885             890             895
Leu Ala Val Leu Gly Pro Gly Tyr Leu Leu Arg Gly Ala Leu Thr Arg

                900             905             910
Val Pro Tyr Phe Val Arg Ala His Ala Leu Leu Arg Met Cys Thr Met
        915             920             925
Val Arg His Leu Ala Gly Gly Arg Tyr Val Gln Met Ala Leu Leu Ala
        930             935             940
Leu Gly Arg Trp Thr Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met
945             950             955                     960
Ser Asp Trp Ala Ala Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu
            965             970             975
Pro Ile Ile Phe Ser Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala
        980             985             990
Glu Thr Ala Ala Cys Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala
        995             1000            1005
Arg Leu Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser
    1010            1015            1020
Lys Gly Trp Lys Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr
1025            1030            1035                    1040
Arg Gly Leu Leu Gly Ser Ile Val Val Ser Met Thr Gly Arg Asp Lys
            1045            1050            1055
Thr Glu Gln Ala Gly Glu Val Gln Val Leu Ser Thr Val Thr Gln Ser
        1060            1065            1070
Phe Leu Gly Thr Ser Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly
        1075            1080            1085
Ala Gly Asn Lys Thr Leu Ala Gly Ser Arg Gly Pro Val Thr Gln Met
    1090  ˏ         1095            1100
Tyr Ser Ser Ala Glu Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly
1105            1110            1115                    1120
Thr Lys Ser Leu Glu Pro Cys Thr Cys Gly Ala Val Asp Leu Tyr Leu
            1125            1130            1135
Val Thr Arg Asn Ala Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys
        1140            1145            1150
```

43

Arg Gly Ala Leu Leu Ser Pro Arg Pro Leu Ser Thr Leu Lys Gly Ser
    1155   1160   1165
Ser Gly Gly Pro Val Leu Cys Pro Arg Gly His Ala Val Gly Ile Phe
   1170   1175   1180
Arg Ala Ala Val Cys Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile
1185   1190   1195   1200
Pro Val Glu Thr Leu Asp Ile Val Thr Arg Ser Pro Thr Phe Ser Asp
                1205   1210   1215
Asn Ser Thr Pro Pro Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu
                1220   1225   1230
His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr
        1235   1240   1245
Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala
   1250   1255   1260
Thr Leu Gly Phe Gly Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro
1265   1270   1275   1280
Asn Ile Arg Thr Gly Val Arg Thr Val Thr Thr Gly Glu Pro Ile Thr
                1285   1290   1295
Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ala Gly Gly
        1300   1305   1310
Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Val Asp Ala Thr
        1315   1320   1325
Thr Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly
   1330   1335   1340
Val Arg Leu Thr Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr
1345   1350   1355   1360
Thr Pro His Pro Asn Ile Glu Glu Val Ala Leu Gly Gln Glu Gly Glu
                1365   1370   1375
Ile Pro Phe Tyr Gly Arg Ala Phe Pro Leu Ser Tyr Ile Lys Gly Gly
        1380   1385   1390
Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala
        1395   1400   1405
Thr Ala Leu Arg Gly Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly
   1410   1415   1420
Leu Asp Val Ser Ile Ile Pro Thr Gln Gly Asp Val Val Val Val Ala
1425   1430   1435   1440
Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile
                1445   1450   1455
Asp Cys Asn Val Ala Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro
        1460   1465   1470
Thr Phe Thr Ile Thr Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg
        1475   1480   1485
Ser Gln Arg Arg Gly Arg Thr Gly Arg Gly Arg Leu Gly Ile Tyr Arg
   1490   1495   1500
Tyr Val Ser Thr Gly Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val
1505   1510   1515   1520
Leu Cys Glu Cys Tyr Asp Ala Gly Ala Ala Trp Tyr Glu Leu Ser Pro
                1525   1530   1535
Val Glu Thr Thr Val Arg Leu Arg Ala Tyr Phe Asn Thr Pro Gly Leu
        1540   1545   1550
Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Ala Val Phe Thr Gly
        1555   1560   1565
Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly
   1570   1575   1580
Glu Asn Phe Ala Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg
1585   1590   1595   1600
Ala Lys Ala Pro Pro Pro Ser Trp Asp Val Met Trp Lys Cys Leu Thr

44

```
                1605                    1610                      1615
Arg Leu Lys Pro Thr Leu Val Gly Pro Thr Pro Leu Leu Tyr Arg Leu
            1620                    1625                    1630
Gly Ser Val Thr Asn Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr
        1635                    1640                    1645
Ile Ala Thr Cys Met Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp
    1650                    1655                    1660
Val Leu Ala Gly Gly Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala
1665                    1670                    1675                    1680
Thr Gly Cys Val Ser Ile Ile Gly Arg Leu His Ile Asn Gln Arg Ala
                1685                    1690                    1695
Val Val Ala Pro Asp Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met
            1700                    1705                    1710
Glu Glu Cys Ala Ser Arg Ala Ala Leu Leu Glu Glu Gly Gln Arg Ile
        1715                    1720                    1725
Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser
    1730                    1735                    1740
Lys Gln Ala Gln Asp Ile Gln Pro Ala Val Gln Ala Ser Trp Pro Lys
1745                    1750                    1755                    1760
Met Glu Gln Phe Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile
            1765                    1770                    1775
Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala
        1780                    1785                    1790
Ser Met Met Ala Phe Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser
    1795                    1800                    1805
Thr Thr Ile Leu Leu Asn Ile Leu Gly Gly Trp Leu Ala Ser Gln Ile
    1810                    1815                    1820
Ala Pro Pro Ala Gly Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly
1825                    1830                    1835                    1840
Ala Ala Val Gly Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu
            1845                    1850                    1855
Ala Gly Tyr Gly Ala Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile
            1860                    1865                    1870
Met Ser Gly Glu Lys Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro
        1875                    1880                    1885
Gly Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala
    1890                    1895                    1900
Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met
1905                    1910                    1915                    1920
Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr
            1925                    1930                    1935
His Tyr Val Thr Glu Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu
        1940                    1945                    1950
Gly Ser Leu Thr Ile Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile
        1955                    1960                    1965
Thr Glu Asp Cys Pro Ile Pro Cys Ala Gly Ser Trp Leu Arg Asp Val
    1970                    1975                    1980
Trp Asp Trp Val Cys Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr
1985                    1990                    1995                    2000
Ser Lys Leu Phe Pro Lys Met Pro Gly Leu Pro Phe Ile Ser Cys Gln
                2005                    2010                    2015
Lys Gly Tyr Lys Gly Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg
            2020                    2025                    2030
Cys Pro Cys Gly Ala Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met
        2035                    2040                    2045
Arg Ile Thr Gly Pro Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe
    2050                    2055                    2060
```

45

Pro Ile Asn Cys Tyr Thr Glu Gly Gln Cys Leu Pro Lys Pro Ala Leu
2065   2070   2075   2080
Asn Phe Lys Thr Ala Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu
                2085   2090   2095
Val Thr Gln His Gly Ser Tyr Ala Tyr Ile Thr Gly Leu Thr Thr Asp
                2100   2105   2110
Asn Leu Lys Val Pro Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp
            2115   2120   2125
Val Asp Gly Val Gln Ile His Arg Ser Ala Pro Thr Pro Lys Pro Phe
            2130   2135   2140
Phe Arg Asp Glu Val Ser Phe Ser Val Gly Leu Asn Ser Phe Val Val
2145   2150   2155   2160
Gly Ser Gln Leu Pro Cys Asp Pro Glu Pro Asp Thr Glu Val Val Met
                2165   2170   2175
Ser Met Leu Thr Asp Pro Ser His Ile Thr Ala Glu Ala Ala Ala Arg
                2180   2185   2190
Arg Leu Ala Arg Gly Ser Pro Pro Ser Glu Ala Ser Ser Ser Ala Ser
            2195   2200   2205
Gln Leu Ser Ala Pro Ser Leu Arg Ala Thr Cys Thr Thr His Gly Arg
            2210   2215   2220
Thr Tyr Asp Val Asp Met Val Asp Ala Asn Leu Phe Met Gly Gly Gly
2225   2230   2235   2240
Val Ile Arg Ile Glu Ser Glu Ser Lys Val Val Val Leu Asp Ser Leu
                2245   2250   2255
Asp Ser Met Thr Glu Glu Glu Gly Asp Leu Glu Pro Ser Val Pro Ser
                2260   2265   2270
Glu Tyr Met Leu Pro Arg Lys Arg Phe Pro Pro Ala Leu Pro Ala Trp
            2275   2280   2285
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Ser Trp Lys Arg Pro
            2290   2295   2300
Asp Tyr Gln Pro Pro Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys
2305   2310   2315   2320
Lys Thr Pro Thr Pro Pro Pro Arg Arg Arg Arg Thr Val Gly Leu Ser
                2325   2330   2335
Glu Ser Thr Ile Gly Asp Ala Leu Gln Gln Leu Ala Ile Lys Ser Phe
                2340   2345   2350
Gly Gln Pro Pro Pro Ser Gly Asp Ser Gly Leu Ser Thr Gly Ala Asp
            2355   2360   2365
Ala Ala Asp Ser Gly Asp Arg Thr Pro Pro Asp Glu Leu Ala Leu Ser
            2370   2375   2380
Glu Thr Gly Ser Thr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly
2385   2390   2395   2400
Asp Pro Asp Leu Glu Pro Glu Gln Val Glu Leu Gln Pro Pro Pro Gln
            2405   2410   2415
Gly Gly Glu Ala Ala Pro Gly Ser Asp Ser Gly Ser Trp Ser Thr Cys
            2420   2425   2430
Ser Glu Glu Asp Asp Ser Val Val Cys Cys Ser Met Ser Tyr Ser Trp
            2435   2440   2445
Thr Gly Ala Leu Ile Thr Pro Cys Ser Pro Glu Glu Lys Leu Pro
            2450   2455   2460
Ile Asn Ser Leu Ser Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr
2465   2470   2475   2480
Cys Thr Thr Ser Lys Ser Ala Ser Leu Arg Ala Lys Lys Val Thr Phe
            2485   2490   2495
Asp Arg Met Gln Val Leu Asp Ala Tyr Tyr Asp Ser Val Leu Lys Asp
            2500   2505   2510
Ile Lys Leu Ala Ala Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu

```
            2515                    2520                    2525
Glu Ala Cys Gln Leu Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly
    2530                    2535                    2540
Phe Gly Ala Lys Glu Val Arg Ser Leu Ser Gly Arg Ala Val Asn His
2545                    2550                    2555                    2560
Ile Lys Ser Val Trp Lys Asp Leu Leu Glu Asp Ser Gln Thr Pro Ile
                2565                    2570                    2575
Pro Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala
            2580                    2585                    2590
Lys Gly Gly Lys Lys Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly
        2595                    2600                    2605
Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Thr Gln Lys Leu
    2610                    2615                    2620
Pro Gln Ala Val Met Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala
2625                    2630                    2635                    2640
Gln Arg Val Glu Phe Leu Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro
                2645                    2650                    2655
Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu
            2660                    2665                    2670
Arg Asp Ile Arg Thr Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro
        2675                    2680                    2685
Glu Glu Ala Arg Thr Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val
    2690                    2695                    2700
Gly Gly Pro Met Phe Asn Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg
2705                    2710                    2715                    2720
Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr
                2725                    2730                    2735
Cys Tyr Val Lys Ala Leu Ala Ala Cys Lys Ala Ala Gly Ile Ile Ala
            2740                    2745                    2750
Pro Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser
        2755                    2760                    2765
Gln Gly Thr Glu Glu Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala
    2770                    2775                    2780
Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr
2785                    2790                    2795                    2800
Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu
                2805                    2810                    2815
Gly Pro Gln Gly Arg Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr
            2820                    2825                    2830
Ser Ile Ala Arg Ala Ala Trp Glu Thr Val Arg His Ser Pro Val Asn
        2835                    2840                    2845
Ser Trp Leu Gly Asn Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg
    2850                    2855                    2860
Met Val Leu Met Thr His Phe Phe Ser Ile Leu Met Ala Gln Asp Thr
2865                    2870                    2875                    2880
Leu Asp Gln Asn Leu Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val
                2885                    2890                    2895
Ser Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp
            2900                    2905                    2910
Ala Phe Ser Leu His Thr Tyr Thr Pro His Glu Leu Thr Arg Val Ala
        2915                    2920                    2925
Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Ala Trp Lys Ser
    2930                    2935                    2940
Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala

2945                    2950                    2955                    2960
Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu
```

47

```
                2965                    2970                    2975
Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp
          2980                    2985                    2990
Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg
        2995                    3000                    3005
Ala Arg Pro Arg Leu Leu Leu Leu Ser Leu Leu Leu Leu Ser Val Gly
    3010                    3015                    3020
Val Gly Leu Phe Leu Leu Pro Ala Arg
3025                    3030


<210> 7
<211> 8024
<212> RNA
<213> Artificial Sequence

<220>
<223> HCV-derived synthetically created replicon RNA rSGREP-JFH1/GND

<400> 7
accugccccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccucccg ggagagccau aguggcugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360
ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420
cgcagguucu ccggccgcuu ggguggagag cuauucggc uaugacuggg cacaacagac 480
aaucggcugc ucugaugccg ccguguuccg cugucagcg caggggcgcc cgguucuuuu 540
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccugccgag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780
ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900
cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960
uggcgaugcc ugcuugccga uaucaugu ggaaaaggc cgcuuuucug gauucaucga 1020
cuguggccgg cuggguugug cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080
ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140
ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200
ccucucccuc cccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260
cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320
aaccuggccc ugucuucuug acgagcauuc cuaggggucu uuccccucuc gccaaaggaa 1380
ugcaaggucu guugaauguc gugaaggaag caguccucu ggaagcuucu ugaagacaaa 1440
caacgucugu agcgacccuu gcaggcagc ggaacccccc accuggcgac aggugccucu 1500
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaaccc cagugccacg 1560
uugugaguug auaguuugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620
ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680
caugcuuuac auguguuuag ucgagguuaa aaaacgucu aggccccccg aaccacgggg 1740
acgugguuuu ccuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800
caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860
caggccgggg aaguccaaau ccuguccaca gucucucagu ccuuccucgg aacaaccauc 1920
ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980
gguccgguca cgcagaugua cucgagugcu gaggggacu ugguaggcug cccagccccc 2040
ccuggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100
cggaacgcug augucaucc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160
```

48

```
ccgagaccca uuucgaccuu gaaggggucc ucggggggc cggugcucug cccuaggggc 2220
cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280
uucaucccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340
acgccaccgg cugugcccca gaccuaucag gucgggacu ugcaugcucc aacuggcagu 2400
ggaaagagca ccaagguccc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460
aaccccucgg uagcugccac ccuggggguu ggggcguacc uauccaaggc acauggcauc 2520
aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguaccc 2580
acauauggca aauuucucgc cgauggggc ugcgcuagcg gcgccaugu caucaucaua 2640
ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700
caagcagaga cagccggggu cagacuaacu gugcuggcua cggccacacc ccccgggguca 2760
gugacaaccc cccaucccga uauagaaaag guaggccucg ggcgggaggg ugagaucccc 2820
uucuauggga gggcgauucc ccuauccugc aucaagggag ggagacaccu gauuuucugc 2880
cacucaaaga aaaagugug cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940
guggcauacu auagagggu ggacgucucc auaauaccag cucagggaga uguggugguc 3000
gucgccaccg acgcccucau gacgggguac acuggagacu uugacuccgu gaucgacugc 3060
aauguagcgg ucacccaagc ugucgacuuc agccuggacc ccaccuucac uauaaccaca 3120
cagacugucc cacaagacgc ugucucacgc agucagcgcc gcgggcgcac agguagagga 3180
agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240
guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300
accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360
cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucucc 3420
caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480
gccagagcca aggccccucc cccguccugg gacgccaugu ggaagugccu ggcccgacuc 3540
aagccuacgc uugcgggccc cacaccucuc cuguaccguu ugggcccuau uaccaaugag 3600
gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660
gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720
cuggcgacug gaugcguuuc caucaucggc cgcuugcacg ucaaccagcg agucgucguu 3780
gcgccggaua aggaggurcu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840
gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900
uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960
cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020
cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080
gccgcccuca ccaguccguu gucgaccagu accaccaucc uucucaacau cauggaggc 4140
ugguuagcgu cccagaucgc accacccgcg ggggccaccg gcuuugucgu cagguggccug 4200
gugggggcug ccguggcag cauaggccug gguaaggugc uggaggacau ccuggcagga 4260
uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320
ucuauggaag augucaucaa ucuacugcc gggauccugu cuccgggagc ccuguggug 4380
ggggucaucu gcgcggccau ucugcgccgc cacguggac cgggggaggg cgcggccaa 4440
uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500
gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560
agccuacuca gaagaccca caauuggaua acgaggacu gccccauccc augcuccgga 4620
uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680
cugaccucua aauuguuccc caagcugccc ggccuccccu ucaucucuug ucaaaagggg 4740
uacaagggug uguggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800
aucucuggca augucgcu gggcucuaug aggaucacag ggccaaaac cugcaugaac 4860
accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcgc gccgaaaccc 4920
cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980
cagcauggu cguacuccua uguaacagga cugaccacug acaaucugaa aauuccuugc 5040
caacuaccuu cuccagaguu uuucuccugg guggacgguu ugcagaucca uagguuugca 5100
cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160
gcugucgggu cccagcuucc cuguguaaccu gagcccgacg cagacguauu gagguccaug 5220
cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280
ccuccaucug aggcgagcuc ucagugagc cagcuaucag caccgucgcu gcgggccacc 5340
ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400
ggcggugugg cucagacaga gccugagucc aggugcccg uucuggacuu ucucgagcca 5460
auggccgagg aagagagcga ccuugagccc ucaauaccau cggagugcau gcuccccagg 5520
agcggguuuc cacgggccuu accggcuugg gcacggccug acuacaaccc gccgcucgug 5580
```

```
gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucuccccccc 5640
cccaagaagg ccccgacgcc uccccaagg agacgccgga caguggg ucu gagcgagagc 5700
accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760
ggugaugcag gcucguccac gggggcgggc gccgccgaau ccggcggucc gacguccccu 5820
ggugagccgg ccccucaga gacagguucc gccuccucua ugcccccccu cgaggggag 5880
ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucccc ccaggggggg 5940
gggguagcuc ccgguucggg cucgggucu uggucuacuu gcuccgagga ggacgauacc 6000
accgugugcu gcuccauguc auacuccugg accggggcuc uaauaaccc cuguagcccc 6060
gaagaggaaa aguugccaau caacccuuug aguaacugc uguugcgaua ccauaacaag 6120
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180
acgcaagugc ucgacgccca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240
aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300
gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360
aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420
accaucaugg ccaaaaauga ggguuucugc guggacccccg ccaagggggg uaagaaacca 6480
gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540
gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600
ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccauGGGu 6660
uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720
gaguccauau accaggccug cucccugccc gaggaggccc gcacugccau acacucgcug 6780
acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840
agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900
gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau gcugguaugc 6960
ggcaaugacc uaguagucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020
agagccuuca cggaggccau gaccagguac ucugccccuc cuggugaucc ccccagaccg 7080
gaauaugacc uggagcuaau aacauccugu uccucaaaug ugucuguggc guugggcccg 7140
cggggccgcc gcagauacua ccugaccaga gacccaacca uccacucgc ccgggcugcc 7200
ugggaaacag uuagacacuc cccuaucaau ucauggcugg gaaacaucau ccaguaugcu 7260
ccaaccauau ggguucgcau gguccuaaug acacacuucu ucuccauucu caugguccaa 7320
gacacccugg accagaaccu caacuuugag auguauggau caguauacuc cgugaauccu 7380
uuggaccuuc cagccauaau ugagagguua cacggcuug acgccuuuuc uaugcacaca 7440
uacucucacc acgaacugac gcgggggcu ucagcccuca gaaaacuugg ggcgccaccc 7500
cucaggGuGu ggaagagucg ggcucgcgca gucagggcgu cccucaucuc ccguggaggg 7560
aaagcggccg uuugcggccg auaucucuuc aauugggcgg ugaagaccaa gcucaaacuc 7620
acuccauugc cggaggcgcg ccuacuggac uuauccaguu gguucaccgu cggcgccggc 7680
gggggcgaca uuuuucacag cgugucgcgc gcccgacccc gcucauuacu cuucggccua 7740
cuccuacuuu ucguagggu aggccucuuc cuacucccg cucgguagag cggcacacac 7800
uagguacacu ccauagcuaa cuguuccuuu uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu 7860
uuuuuuuuuu cuuuuuuuuu uuuuucccuc uuucuucccu ucucaucuua uucuacuuuc 7920
uuucuuggug gcuccaucuu agcccagucc acggcuagcu gugaaagguc cgugagccgc 7980
augacugcag agagugccgu aacuggucuc ucugcagauc augu            8024
```

```
<400> 8
accugccccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucuacagc cuccaggccc 120
cccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
```

50

```
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360
ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420
cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480
aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggcgcc cgguucuuuu 540
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccugccgag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780
ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucagggc ucgcgccagc 900
cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960
uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020
cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080
ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140
ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200
ccucucccuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260
cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320
aaccuggccc ugucuucuug acgagcauuc cuaggggucu uucccoucuc gccaaaggaa 1380
ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440
caacgucugu agcgacccuu ugcaggcagc ggaaccccc accggcgac aggugccucu 1500
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaaccc cagugccacg 1560
uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620
ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680
caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggccccccg aaccacgggg 1740
acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800
caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860
caggccgggg aaguccaaau ccuguccaca gucucucagu ccuuccucgg aacaaccauc 1920
ucgggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980
gguccgguca cgcagaugua cucgagugcu gaggggacu ugguaggcug gcccagcccc 2040
ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100
cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160
ccgagaccca uuucgaccuu gaaggggucc ucggggggggc cggugcucug cccuaggggc 2220
cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280
uucauccccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340
acgccaccgg cugugcccca gaccaucag gucggguacu ugcaugcucc aacuggcagu 2400
ggaaagagca ccaagguccc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460
aaccccucgg uagcugccac ccuggggUuu ggggcguacc uauccaaggc acauggcauc 2520
aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580
acauauggca aauuucucgc cgauggggc ugcgcuagcg gcgccuauga caucaucaua 2640
ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700
caagcagaga cagccggggu cagacuaacu gugcuggcua cggccacacc ccccgggguca 2760
gugacaaccc cccaucccga uauagaagag guaggccucg ggcgggaggg ugagauccc 2820
uucuauggga gggcgauucc ccuauccugc aucaagggag ggagacaccu gauuuucugc 2880
cacucaaaga aaaagugga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940
guggcauacu auagagggUu ggacgucucc auaauaccag cucaggagga uguggugguc 3000
gucgccaccg acgcccucau gacggggUac acuggagacu uugacuccgu gaucgacugc 3060
aauguagcgg ucacccaagc ugucgacuuc agccuggacc ccaccuucac uauaaccaca 3120
cagacugucc cacaagacgc ugucucacgc agucagcgcc gcgggcgcac agguagagga 3180
agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240
guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300
accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360
cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucucc 3420
caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480
gccagagcca aggcccuucc cccguccugg gacgccaugu ggaagugccu ggcccgacuc 3540
aagccuacgc uugcggggcccc cacaccucuc cuguaccguu ugggcccuau uaccaaugag 3600
gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660
gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720
```

```
cuggcgacug gaugcguuuc caucaucggc cgcuugcacg ucaaccagcg agucgucguu 3780
gcgccggaua aggagguccu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840
gcggcucuca ucgaagaggg gcagcggaua gccgagaugu ugaaguccaa gauccaaggc 3900
uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960
cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020
cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080
gccgcccuca ccaguccguu gucgaccagu accaccaucc uucucaacau caugggaggc 4140
ugguuagcgu cccagaucgc accacccgcg ggggccaccg gcuuugucgu caguggccug 4200
gugggggcug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260
uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagccc 4320
ucuauggaag augucaucaa ucuacugccu gggauccugu cuccggagc ccugguggug 4380
ggggucaucu gcgcggccau ucugcgccgc cacgugggac cggggagg cgcgguccaa 4440
uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500
gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560
agccuacuca gaagacucca caauuggaua acugaggacu gccccauccc augcuccgga 4620
uccuggcucc gcgacgugug ggacugggu ugcaccaucu ugacagacuu caaaaauugg 4680
cugaccucua aauuguuccc caagcugccc ggccuccccu ucaucucuug ucaaaagggg 4740
uacaagggug uguggccggg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800
aucucuggca auguccgccu gggcucuaug aggaucacag ggccaaaac cugcaugaac 4860
accuggcagg ggaccuuucc uaucaauugc uacacggagg gccagugcgc gccgaaaccc 4920
cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980
cagcaugggu cguacuccua guaacagga cugaccacug acaaucugaa aauuccuugc 5040
caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100
cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160
gcugucgggu cccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220
cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280
ccuccaucug aggcgagcuc cucagugagc cagcuaucag caccgucgcu gcgggccacc 5340
ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400
ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460
auggccgagg aagagagcga ccuugagccc ucaauaccau cggagugcau gcuccccagg 5520
agcggguuuc cacgggccuu accggcuugg gcacggccug acuacaaccc gccgcucgug 5580
gaaucguggaa ggaggccaga uuaccaaccg cccaccguug cugguugugc ucuccccccc 5640
cccaagaagg ccccgacgcc ucccccaagg agacgccgga caguggucu gagcgagagc 5700
accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760
ggugaugcag gcucguccac ggggcgggc gccgccgaau ccggcgguccu gacguccccu 5820
ggugagccgg cccccucaga gacagguuccc gccuccucua ugcccccccu cgaggggag 5880
ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucccccc caggggggg 5940
ggguagcuc ccgguucggg cucgggucu uggucuacuu gcuccgagga ggacgauacc 6000
accgugugcu gcuccaugucc auacuccugg accggggcuc uaauaaacucc cuguagcccc 6060
gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180
acgcaagugc ucgacgccca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240
aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300
gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360
aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauuccacaa 6420
accaucaugg ccaaaaauga ggguucugc guggacccg ccaagggggg uaagaaacca 6480
gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540
gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600
ccugcccaac gggugguagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660
uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720
gaguccauau accaggccug cucccugccc gaggaggccc gcacugccau acacucgcug 6780
acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840
agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900
gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960
caggggacug aggaggacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020
ucugccccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauuccgu 7080
uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140
```

52

```
gacccaacca cuccacucgc ccgggcugcc ugggaaacag uuagacacuc cccuaucaau 7200
ucauggcugg gaaacaucau ccaguaugcu ccaaccauau ggguucgcau gguccuaaug 7260
acacacuucu ucuccauucu caugguccaa gacacccugg accagaaccu caacuuugag 7320
auguauggau caguauacuc cgugaauccu uggaccuuc cagccauaau ugagagguua 7380
cacgggcuug acgccuuuuc uaugcacaca uacucucacc acgaacugac gcgggguggcu 7440
ucagcccuca gaaaacuugg ggcgccaccc cucaggggugu ggaagagucg ggcucgcgca 7500
gucagggcgu cccucaucuc ccguggaggg aaagcggccg uuugcggccg auaucucuuc 7560
aauugggcgg ugaagaccaa gcucaaacuc acuccauugc cggaggcgcg ccuacuggac 7620
uuauccaguu gguucaccgu cggcgccggc gggggcgaca uuuuucacag cgugucgcgc 7680
gcccgacccc gcucauuacu cuucggccua cuccuacuuu ucguaggggu aggccucuuc 7740
cuacuccccg cucgguagag cggcacacac uagguacacu ccauagcuaa cuguuccuuu 7800
uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu cuuuuuuuuu uuuuucccuc 7860
uuucuucccu ucucaucuua uucuacuuuc uuucuugguag gcuccaucuu agcccuaguc 7920
acggcuagcu gugaaagguc cgugagccgc augacugcag agagugccgu aacggucuc 7980
ucugcagauc augu 7994


<210> 9
<211> 340
<212> RNA
<213> Artificial Sequence


<220>
<223> Description of Artificial Sequence: synthetic RNA


<400> 9
accugccccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
cgcuugcgag ugccccggga ggucucguag accgugcacc 340


<210> 10
<211> 340
<212> RNA
<213> Artificial Sequence


<220>
<223> Description of Artificial Sequence: synthetic RNA


<400> 10
acccgccccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60
cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
ugcuugcgag ugccccggga ggucucguag accgugcacc 340


<210> 11
<211> 236
<212> RNA
<213> Artificial Sequence


<220>
```

53

<400> 11
agcggcacac acuagguaca cuccauagcu aacuguuccu uuuuuuuuuu uuuuuuuuuu 60
uuuuuuuuuu uuuuuuuuuu uucuuuuuuu uuuuuuuccc ucuuucuucc cuucucaucu 120
uauucuacuu ucuuucuugg uggcuccauc uuagcccuag ucacggcuag cugugaaagg 180
uccgugagcc gcaugacugc agagagugcc guaacugguc ucucugcaga ucaugu   236

<210> 12
<211> 232
<212> RNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic RNA

<400> 12
agcggcacac auuagcuaca cuccauagcu aacuguuccu uuuuuuuuuu uuuuuuuuuu 60
uuuuuuuuuu uuuuuuucuu uuuuuuuuuu uuucccucuu ucuucccuuc ucaucuuauu 120
cuacuuucuu ucuugguggc uccaucuuag cccuggucac ggcuagcugu gaaagguccg 180
ugagccgcau gacugcagag agugccguaa cuggucucuc ugcagaucau gu   232

<210> 13
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 13
cgggagagcc atagtgg   17

<210> 14
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 14
agtaccacaa ggcctttcg   19

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 15

```
ctgcggaacc ggtgagtaca c                                              21

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 16
aacaagatgg attgcacgca                                                20


<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 17
cgtcaagaag gcgatagaag                                                20


<210> 18
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 18
gcactctctg cagtcatgcg gctcacggac                                     30


<210> 19
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 19
cccctgtgag gaactactgt cttcacgc                                       28


<210> 20
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA
```

<400> 20
ccgggagagc catagtggtc tgcg   24

<210> 21
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 21
ccactcaaag aaaaagtgtg acgagctcgc   30

<210> 22
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 22
ggcttgggca cggcctga   18

<210> 23
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 23
gcggtgaaga ccaagctcaa actcactcca   30

<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 24
agaacctgcg tgcaatccat c   21

<210> 25
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 25
cccgtcatga gggcgtcggt ggc   23

<210> 26
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 26
accagcaacg gtgggcggtt ggtaatc   27

<210> 27
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 27
ggcacgcgac acgctgtg   18

<210> 28
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 28
agctagccgt gactagggct aagatggagc   30

<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      DNA(primer)

<400> 29
aacaagatgg attgcacgca   20

```
<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
      DNA(primer)

<400> 30
cgtcaagaag gcgatagaag                                                20


<210> 31
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 31
gcactctctg cagtcatgcg gctcacggac                                     30


<210> 32
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 32
cccctgtgag gaactactgt cttcacgc                                       28


<210> 33
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 33
ccgggagagc catagtggtc tgcg                                           24


<210> 34
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA
```

<400> 34
ccactcaaag aaaaagtgtg acgagctcgc   30

<210> 35
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
     DNA(primer)

<400> 35
ggcttgggca cggcctga   18

<210> 36
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 36
gcggtgaaga ccaagctcaa actcactcca   30

<210> 37
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 37
agaacctgcg tgcaatccat c   21

<210> 38
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 38
cccgtcatga gggcgtcggt ggc   23

<210> 39
<211> 27
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 39
accagcaacg gtgggcggtt ggtaatc   27

<210> 40
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 40
ggaacgcgac acgctgtg   18

<210> 41
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic DNA

<400> 41
agctagccgt gactagggct aagatggagc   30